Humboldt-Universität zu Berlin

Dissertation

Extracting muscle synergies from human steady and unsteady locomotion: methods and experiments

Dissertation

zur Erlangung des akademischen Grades Dr. phil.

im Fach Sportwissenschaft

Kultur-, Sozial- und Bildungswissenschaftliche Fakultät der Humboldt-Universität zu Berlin

Alessandro Santuz

Dekan: Prof. Dr. Julia von Blumenthal

Gutachter/in: 1. Prof. Kiros Karamanidis

2. Prof. Hendrik Schmidt

Datum der Promotion: 02.05.2018

Zusammenfassung

Die Notwendigkeit, sich über unebene, sich ständig verändernde Gelände und Umgebungen zu bewegen, gehört zu unserem täglichen Leben. Die ununterbrochene Quelle von Störungen, die durch Veränderungen der Morphologie des Bodens erzeugt werden, ist Ursache für eine ständige Instabilität. Das zentrale Nervensystem muss daher eine erhöhte Menge an Information integrieren, um mit der Unvorhersehbarkeit äußerer Störungen zurechtkommen zu können. Die Folge dieser erhöhten Beanspruchung könnte eine flexible Kombination der modularen Organisation von Bewegungsentstehung und -steuerung sein. Auf Kosten der Genauigkeit der Bewegung wäre es so möglich, dass das System reagiert, indem es die Robustheit seiner Steuerung erhöht (d.h. die Fähigkeit mit Fehlern oder stressigen Umgebungsbedingungen umzugehen). Jedoch sind die neuronalen Strategien, die das zentrale Nervensystem zur Organisation der Bewegung verwendet, immer noch schlecht verstanden. Es existieren jedoch mehrere Theorien. Eine Möglichkeit besteht darin, dass Bewegungen nicht durch die kostenintensive Kontrolle mehrerer Muskeln unabhängig voneinander zustande kommen, sondern durch eine kleine Anzahl linear kombinierter Aktivierungsmuster, die Muskelsynergien genannt werden. Die komplexen elektromyographischen Pattern, die in den Muskeln gelesen werden können, könnten dann von wenigen niederdimensionalen Einheiten über efferente oder afferente Wege erzeugt werden.

Unter den verschiedenen Möglichkeiten der Bewegungsstörung sind das Weglassen von Schuhen und die Verwendung von unebenen Oberflächen zwei gebräuchliche Optionen. In einem ersten Schritt habe ich eine gründliche Analyse der Methoden durchgeführt, die nützlich sind für a) die Auswertung von raumzeitlichen Gangparametern mithilfe von Daten der plantaren Druckverteilung (erste Studie vorgestellt in dieser Arbeit) und b) die Extraktion von Muskelsynergien mittels nicht-negativer Matrixfaktorisierung (zweite Studie). Anschließend habe ich die modulare Organisation von c) beschut und barfuß Laufen (dritte Studie) und d) Laufband Gehen und Laufen über ebener und unebener Oberfläche (vierte Studie) analysiert. Die modulare Organisation der Fortbewegung, bewertet durch die Extraktion von Muskel-Synergien, änderte sich, als Störungen eingeführt wurden. Im Vergleich zum gestörten Zustand zeigte das Barfußlaufen eine zeitlichen Verschiebung der zeitabhängigen Muskelaktivierungspatterns (Motor Primitives) und eine Reorganisation der zeitunabhängigen Koeffizienten (Motor Modules), hauptsächlich aufgrund des unterschiedlichen Fußaufsatzes. Zusammenfassend, konserviert Fortbewegung über unebener Oberfläche, im Vergleich zu ebener, Motor Modules, während Motor Primitives im Allgemeinen breiter werden. Diese Ergebnisse unterstützen die Idee einer erhöhten Robustheit in der motorischen Kontrolle während der instabilen Fortbewegung.

Schlagwörter:

Muscle Synergies, Motor Control, Human Locomotion, Biomechanics, Neuroscience

Abstract

The need to move over uneven, continuously changing terrains and environments is part of our daily life. The uninterrupted source of perturbations generated by alterations of ground's morphology, is cause of relentless instability. Thus, the central nervous system must integrate an augmented amount of information in order to be able to cope with the unpredictability of external disturbances. A consequence of this increased demand might be a flexible recombination of the modular organisation of movement creation and control. At the expense of motion's accuracy, it is possible that the system responds by increasing its control's robustness (i.e. ability to cope with errors or stressful environmental conditions). However, the neural strategies employed by the central nervous system to organise movement are still poorly understood, although several theories exist. One possibility is that movements are constructed not by the costly control of several muscles independently, but through a small amount of linearly combined patterns of activations, called muscle synergies. The complex electromyographic patterns readable in muscles, could be then produced by just a few low dimensional units via efferent or afferent pathways.

Amongst the several possibilities of perturbing locomotion, the removal of footwear and the use of uneven surfaces are two valid options. In a first step, I conducted a thorough analysis of the methodologies useful for a) the evaluation of spatiotemporal gait parameters using plantar pressure distribution data (first study presented in this thesis) and b) the extraction of muscle synergies using non-negative matrix factorisation (second study). Afterwards, I analysed the modular organisation of c) shod and barefoot running (third study) and d) walking and running over an even- and an uneven-surface treadmill (fourth study). The modular organisation of locomotion, assessed through the extraction of muscle synergies, changed when perturbations were introduced. Compared to the shod condition, barefoot running underwent, mostly due to the different foot strike pattern, a reorganisation of the time-independent coefficients (motor modules) and a time-shift of the time-dependent muscle activation patterns (motor primitives). Uneven-surface locomotion, compared to even-surface, conserved motor modules, while motor primitives were generally wider, confirming the idea of an increased robustness in motor control during unsteady locomotion.

Keywords:

Muscle Synergies, Motor Control, Human Locomotion, Biomechanics, Neuroscience

Table of Content

To my family

in gratitude and love

1 Introduction

In the beginning of the XIX century, the French physician and neurophysiologist Marie Jean Pierre Flourens (1794 – 1867) wrote (freely translated from French):

> "So there are, in the nervous system, three essentially distinct properties:
> One of *willing* and *sensing*; it is the *perception*;
> The other of *exciting instantly the muscle contraction*; I suggest to call it *excitability*;
> The third of *coordinating* the movements; I call it *coordination*."[1]

> "Il y a donc, dans le système nerveux, trois propriétés essentiellement distinctes:
> L'une de *vouloir* et de *percevoir*; c'est la *sensibilité*;
> L'autre d'*exciter immédiatement la contraction musculaire*; je propose de l'appeler *excitabilité*;
> La troisième de *coordonner* le mouvemens; je l'appelle *coordination*."[1]

Almost 200 years later, how the central nervous system (CNS) copes with the complexity of controlling multiple joints and muscles for producing and controlling movement is still an open question in neuroscience.

1.1 Central Pattern Generators

The CNS is composed by the brain and the spinal cord. However, a considerable portion of movement production is related to relatively simple, organised spinal neuronal pathways, the so-called central pattern generators (CPGs)[2–8]. With their ability to produce rhythmic, supraspinally-independent movement patterns shaped by sensory and neuromodulatory inputs[2–5,7], the CPGs play a major role in the motor control of those cyclic activities such as, for instance, locomotion (e.g. walking, running, swimming, etc.)[2–8]. Locomotion is a repetitive, highly-stereotyped movement; thus, it represents an ideal object for the study of movement creation and control in a scientific experimental setup[9]. In order to understand the structure and organisation of CPGs for locomotion, the anatomy and function of the spinal cord must be first briefly presented.

1.1.1 Gross anatomy and function of the spinal cord

The spinal cord is the most caudal portion of the CNS. In adult humans, it extends from the lowest region of the brain (the medulla oblongata) to the lumbar segments of the vertebral column, which encloses it in a strong and flexible protective structure. Its length is around 43 to 45 cm from the medulla to the first lumbar vertebra and it has an elliptical cross-section[2]. The spinal cord is divided into segments. In humans, there are generally 31 segments, but this number is species-specific[2]. Each spinal segment gives rise to paired spinal nerves (i.e. bundles of neuronal axons) joined in dorsal and ventral roots (see Figure 1 for a graphical representation). A dorsal root contains afferent fascicles and is characterised by the distal presence of spinal ganglia, containing the somas (i.e. the bodies) of the neuronal cells. A ventral root consists of the axons of those motor neurons that have their bodies in the ventral (or anterior) grey horns of the spinal cord. The interneurons relay and modulate the information between the dorsal and ventral roots. The human spinal cord is made of 31 nerve segments: eight cervical segments forming eight pairs of cervical nerves (C1 to C8); twelve thoracic segments forming twelve pairs of thoracic nerves (T1 to T12); five lumbar segments forming five pairs of lumbar nerves (L1 to L5); five sacral segments forming five pairs of sacral nerves (S1 to S5); one coccygeal segment. Cervical segments accommodate those innervations useful to control muscles involved in respiration and in the movement of head, neck and arms[2]. Thoracic segments are important in controlling the movement of fingers and trunk[2]. Lastly, the lumbar and sacral segments are connected to the pathways for controlling locomotion, urination, intestinal and reproductive functions[2].

In the spinal cord both white and grey matter tissues are present. The former contains ascending (to the brain) and descending (to the periphery) myelinated fibres, including sensory and motor neuronal axons. The latter encloses unmyelinated neuronal cells (called interneurons) as well as the somas, axons and dendrites of motor neurons, embedding simple (e.g. reflex arcs) and more complex (e.g. CPGs) spinal circuits. Simple, locally-mediated reflex pathways such as the Ia, Ib and FRA (flexion reflex afferent) reflex arcs are, in order of decreasing simplicity as to the number of synaptic connections, well described in literature[2]. The excitatory Ia (monosynaptic) and Ib (disynaptic) reflex arcs are thought to be major contributors for muscular contraction during locomotion by mediating afferent inputs from muscle spindles and Golgi tendon organs (GTOs), respectively[2,10,11]. The FRA, even if

typically involved in limb's withdrawal when receiving information from cutaneous nociceptors (receptors of potential harm), is also an important reflex for locomotion. In fact, the FRA can enhance postural control and has been shown, if pharmacologically stimulated, to reset the gait cycle to flexion[10,11].

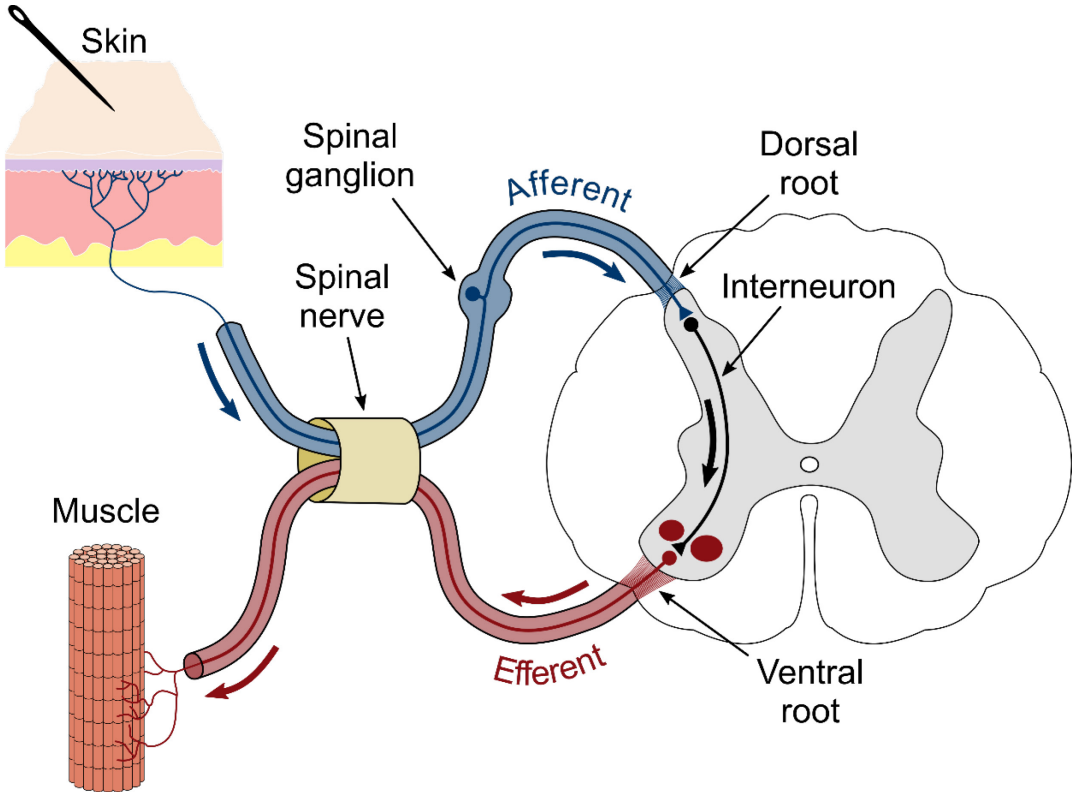

Figure 1 Schematic representation of a spinal reflex arc. A pin in the skin produces an input signal which travels through the afferent pathway in the spinal nerve until the dorsal root of the spinal cord. The relevant motor output is exiting the spinal cord from the ventral root and finally reaches the muscle. The connection between the afferent (input) and efferent (output) pathways is mediated by the interneuron.

Several white matter tracts have the function of relaying information between the brain and the spinal cord. Ascending pathways that travel to the brain guide sensory-related signals such as touch, proprioception, pressure and vibration. These large myelinated fibres decussate (travel contralaterally) through the spinal cord until they reach the upper levels of the thalamus and sensory cortex. Descending pathways that travel from the brain are the corticospinal (or pyramidal) and the extrapyramidal tracts. These pathways carry the information associated with complex movements, such as fine skilled motion, dynamic stability maintenance and stereotyped motor behaviours like locomotion.

The scientific attention given to the cellular organisation and function of the spinal grey matter is relatively recent. On the one hand, somatic (i.e. related to skeletal muscle) reflex arcs have been known and studied for almost two centuries[1,12]. In his book published in 1824, Flourens described how (freely translated from French) "transecting the lumbar region [of the spinal cord in different animals], all the muscles innervated by the nerves of this region moved together [but not] spontaneously nor voluntarily."[1]. Just fifty years after Flourens, Freusberg reported that (freely translated from German) "we saw […] dogs standing on four legs and walk [after a complete transection of the spinal cord], […] just days after surgery"[13]. On the other hand, more complex spinal networks such as the CPG for locomotion have been intensively analysed only with the advent of intracellular recordings, in the late 1960s[2].

1.1.2 The CPG for locomotion

In the first years of the 20th century, Sir Charles Scott Sherrington (1932 Nobel Prize in Physiology or Medicine together with Edgar Douglas Adrian "for their discoveries regarding the functions of neurons") conducted an extensive study of spinal, decerebrate and decapitate preparations, mainly in dogs and cats[12,14]. The evidence provided by Sherrington was supporting the idea that locomotion is the result of reflex projections from proprioceptors (receptors of an individual's own body position) onto the spinal cord[12,14]. Moreover, thanks to his experiments involving electrical stimulation of the spinal cord, Sherrington hypothesised the existence of some specialised spinal neurons[12,14], which we call today interneurons. Stepping could be produced by applying a constant flow of current to the exposed cross-section of the cord and would happen in a rhythmic fashion[12,14]. However, the potential existence of an independent spinal neuronal network for locomotion was formally hypothesised by one of Sherrington's junior collaborators[6], Thomas Graham Brown[15]. Graham Brown hypothesised the existence of some "half-centres", spinal entities capable of producing flexion and extension in the absence of descending and/or sensory inputs[6]. Graham Brown argued that the half-centres, in a mutual inhibition fashion, could build rhythmic movements under the important modulation of proprioceptive input[6], an idea he admittedly took and elaborated from Sherrington after he abandoned it. Later, Graham brown further developed this brilliant idea, even including interneuronal spinal connections in his explanations[6]. Nevertheless, the idea that the rhythmic nature of stepping was an intrinsic property of the spinal cord[15] remained controversial for many years.

The term “central pattern generator” first appeared in the 1960s[6,16,17], when the lemma “half-centre”[15] proposed by Graham Brown started being slowly superseded. In his 1914 paper, Graham Brown wrote[15]: “[…] the view here advanced is that the functional unit in the nervous system is not the activity of the reflex arc as is usually assumed, but that it is the activity of a unit which we may term the "half-centre" -or, perhaps, the centre composed of linked antagonistic "half-centres." There remains the question of the manner in which the peripheral reflex should be considered.”. In his seminal observations, Graham Brown was attributing the origin of locomotor activity to the interaction of the flexor and extensor half-centres[15]. Nowadays, the CPGs are described as an ensemble of spinal interneurons and motor neurons whose interplay can produce a variety of rhythmic movements useful for repetitive, highly-stereotyped motor tasks such as walking, running or swimming[2–7]. Given the obvious difficulties to directly examine humans, in the past 70 years several non-human invertebrate species have been used as simple models for studying the CPGs[17,18,27–33,19–26]. Locusts[17,20,21], stick insects[22,23], sea slugs[24,25], crustaceans[26–29], moths[30], fruit flies[31–33] and cockroaches[34,35] have been successfully employed to provide direct evidence for the existence of CPGs. In vertebrates, mainly due to the higher amount of neurons composing the CNS, the efforts to understand the neuronal organisation have started later. However, to date, a fair amount of literature about CPGs in non-mammalian and mammalian vertebrates is available[10,11,42–50,19,25,36–41]. Some of the vertebrate models that received more scientific attention in the past few decades are the cat[10,11,37–40], rat and mouse[41,42], frog[25,38,43–45], turtle[38,46,47], shark[38,48] and lamprey[49,50].

The first experimental proof of the CPGs’ existence was provided by the group of Lundberg in the late 1960s[10,11]. Using L-DOPA injections to stimulate the FRA in spinal cats, Lundberg and his colleagues could show the interplay of ipsi- and contralateral interneurons (FRA and Ia) for producing locomotor patterns[10,11]. These experiments confirmed the modern categorisation of CPGs as a group of interneurons that are mostly located in the lumbar portion of the spinal cord[2]. Since then, however, the organisation of CPGs and their relation with the supraspinal share of motor control have been conceptualised in a multitude of ways. One of these is the muscle synergies model that originated from the theoretical and experimental work of Emilio Bizzi’s group[51,52]. In the next paragraph, the concept of muscle synergies will be further detailed.

1.2 Muscle synergies

The human body is an outstandingly complicated machine. Summing up all the muscles and joints, the total is more or less 1000. Yet, the job of mastering the resulting enormous amount of degrees of freedom for accomplishing movement is left to the only one brain (and spinal cord) we have in our own body. So how is neuromuscular coordination possible? How can we achieve accurate, rapid movements having an outstanding amount of muscle/joint configurations to choose from? In nature, many models exist of discrete systems that can generate meaningful entities just by following some rules[52]. Languages, the genetic code or even more simply those cooking recipes that are part of our daily life are good examples. The text of this thesis is being written by pooling together some symbols, which we call graphemes or letters, and giving them meaning by using the rules of semantics, syntax, etc. What if the CNS was coding movement the same way? What if, in order to overcome the issue of choosing over an excess of degrees of freedom, the CNS was using a combination of discrete elements and their associated rules to produce movement?

1.2.1 Historical note and theory

The etymology of the word "synergy" is nested in the Greek language. Literally, synergy means "working together" (συνεργός). The idea that some synergistic neural components of movement exist was already suggested by Sherrington at the beginning of the XX century. In his famous "The integrative action of the nervous system", Sherrington wrote "The stimulation […] excites reflexly through the central organ an effect in the skeletal musculature which is co-ordinate and synergic."[12]. Yet, Sherrington took some distance from the concept of a functional organisation of the motor spinal root, arguing that "the collection of fibres in a spinal motor root is not a functional collection in the sense that it is representative of any co-ordination"[12].

In 1967, Nikolai Aleksandrovich Bernstein published his "The co-ordination and regulation of movements"[9], a book that became a milestone in the history of muscle synergies. For the first time, Bernstein formally described the so-called "degrees of freedom problem", stating that "the basic difficulties for co-ordination consist precisely in the extreme abundance of degrees of freedom, with which the [CNS] […] is not at first in a position to deal."[9]. This

concept of motor abundance is still one of the supporting pillars of modern motor control and laid the foundation of the muscle synergies idea. Even if Bernstein did not use the lemma "muscle synergies", throughout the text one can read about the concept between the lines. Bernstein wrote that "the co-ordination is certainly not organized independently at the periphery alone and […] is […] centrally determined […] by means of a proprioceptive reflex cycle. The co-ordinational process […] is obviously not a tetanic process, but it undoubtedly incorporates both receptor and effector components."[9]. Following on this path, he added that "locomotor movements display an extremely widespread synergy incorporating the whole musculature and the entire moving skeleton and bringing into play a large number of areas and conduction pathways of the central nervous system."[9]. Moreover, he already gave a very modern definition of coordination: "The co-ordination of a movement is the process of mastering redundant degrees of freedom of the moving organ, in other words its conversion to a controllable system. More briefly, co-ordination is the organization of the control of the motor apparatus."[9]. Bernstein also extensively discussed the concept of motor programs, intended as "a kind of extrapolation between the actually existing situation and what the latter has to become in the interests of the subject."[9,53].

Following the thoughts of Bernstein, in 1984 Lee published an essay in which the idea of "neuromotor synergies", defined as neurally-based units of voluntary action, was explored and supported[54]. The main hypothesis was that "low-level, neurally based patterns significantly constrain intentional actions"[54]. Lee stated that a wide range of voluntary movements can be generated by the interplay of a small set of neural synergies[54]. Also, Lee acknowledged the inherent variability of neuromuscular activity and proposed that invariance in synergies should be determined using statistical approaches[54]. This was an extreme evolution in the field, but it was not until the late 1990s that the adequate statistical tools could be available to transform the purely theoretical ideas of Lee into practice. However, already in 1991, Emilio Bizzi and colleagues were the first to experimentally prove the existence of spinal synergies, which they represented as force fields[51]. It was exactly this idea that eventually led to the modern formulation of the muscle synergies concept.

In the February of 1999, a paper from Bizzi and colleagues put, for the first time, Lee's ideas into numbers[55]. Using a factorisation method based on the non-negative least squares

approach, Bizzi, Tresch and Saltiel could finally give a numerical representation of synergies in the spinal frog[55]. The motor responses were obtained by stimulating several skin locations of the hindlimb. The results showed the movement-specific recruitment of a small set of synergistic muscles, needed to accomplish each response[55].

In the October of the same year, Daniel D. Lee and Sebastian H. Seung published a study that revolutionised the computational approach to extract muscle synergies[56]. By introducing the non-negative matrix factorisation (NMF) to the scientific world, Lee and Seung provided a computational tool for easily extracting synergies from any kind of analysed variable[56]. This second breakthrough, besides making 1999 an outstanding year for the field, opened the door to the exponential growth of the muscle synergies concept in the next two decades, until the present day[56]. Since then, the concept of muscle synergies hit fertile ground in the field of motor control. By recording the electromyographic (EMG) activity of several muscles simultaneously, many research groups could provide indirect support for the existence of a modular organisation of many movement tasks across many species[57]. Although the history of numerical applications is relatively short, a number of review articles on muscle synergies is already available in literature[52,53,65–72,57–64].

1.2.2 Numerical approaches

Muscle synergies are usually extracted from EMG signals. It is possible to build a model based on the linear combination of synergies as basis vectors in the space of muscle activations following the rule

$$\mathbf{m}(t) = \sum_{i=1}^{r} c_i(t)\mathbf{w}_i, \quad c_i(t), \mathbf{w}_i \geq 0 \qquad (1)$$

where $\mathbf{m}(t)$ is a vector containing the time-dependant activations of the recorded muscles at a specific time point t, r is the number of synergies, $\mathbf{w}_i$ is a time-independent vector of non-negative weights and $c_i(t)$ is a time-dependent set of non-negative coefficients[55,57]. The general idea behind this model is that high-dimensional data can be compactly represented by the linear combination of low-dimensional elements. Several methods have been proposed for satisfying this dimensionality reduction needs. Some of them are the principal

component analysis, factor analysis, independent component analysis, vector quantisation, the afore-mentioned NMF, etc. EMG data contains signal-dependent noise, a typical characteristics of those neural structures involved in motor planning and control[73]. The comparison between different factorisation methods applied to EMG data shows the general capability of extracting synergies in a very similar way[56,74]. However, the nature of muscle activation is undeniably non-negative. This is one of the reasons why NMF was and still is one of the most appealing amongst all the factorisation algorithms, given that it does not allow negative entries in the factors[56]. By using matrix notation and the historical NMF naming, one can write Equation 1 as

$$V(t) \approx V_R(t) = WH(t) \quad (2)$$

where V_R represents the reconstructed matrix, which approximates the original matrix V (containing the EMG data)[56]. W is a matrix with dimensions $m \times r$, while H has dimensions $r \times n$, with m being the number of muscles measured, n the number of recorded time points and r the rank of the factorisation (i.e. the number of synergies or maximum number of linearly independent combinations of the base vectors). In general[56], the factorisation rank r is chosen so that $(n + m)r < nm$. The matrix W, which we called the motor modules[75] matrix, contains the time-invariant muscle weightings. The matrix H, which we called the motor primitives[75] matrix, contains the time-dependent coefficients of the factorisation.

Several update rules have been and are continuously proposed for data factorisation via NMF. The classical approach by Lee and Seung is based on maximising the Gaussian likelihood of reconstructing the original EMG signal[56,75–78]. By applying the Equations 3 and 4 to an EMG data set V, one can reconstruct the original data with a certain amount of approximation starting from completely random values of H and W. In Equations 3 and 4, the iteration number is indicated with i, while the capital letter T indicates a transposed matrix (which is the original matrix with row and column indices switched):

$$\begin{cases} H_{i+1} = H_i \dfrac{W_i^T V}{W_i^T W_i H_i} & (3) \\ W_{i+1} = W_i \dfrac{V(H_{i+1})^T}{W_i H_{i+1}(H_{i+1})^T} & (4) \end{cases}$$

Other formulations of the update rules are of course available. Just to give an example, Devarajan and Cheung proposed a model based on a special case of the generalized Inverse-Gaussian distribution[78]. The update rules of this formulation are the following (sub- and superscript notations as in Equations 3 and 4):

$$\begin{cases} H_{i+1} = H_i \sqrt{\dfrac{W_i^T \dfrac{V}{(W_i H_i)^2}}{\dfrac{W_i^T}{V}}} & (5) \\ W_{i+1} = W_i \sqrt{\dfrac{V \dfrac{(H_{i+1})^T}{(W_i H_{i+1})^2}}{\dfrac{(H_{i+1})^T}{V}}} & (6) \end{cases}$$

Different update rules have been shown to provide slightly different results, especially in the dimensionality reduction and reconstruction capabilities[75].

A typical procedure for convergence and choice of the rank r (i.e. the minimum number of synergies necessary to reconstruct the original data set) is the following. The quality of the reconstruction V_R of V is calculated with the coefficient of determination R^2. When a change in the calculated R^2 between V and V_R is smaller than the 0.01% in the last 20 iterations[75,79], one can set the limit of convergence. This is done for a number of synergies successively increased from 1 to a number smaller than the dimension m of V, or until the following is satisfied: $(n + m)r < nm$. The computation is repeated several times (10 or 20 are common values) for each synergy, each time creating new randomised initial matrices H and W, in order to avoid local minima[75,80]. The coefficient of determination R^2, expressed by (1-*RSS/SST*), where RSS is the residual sum of squares and SST is the total sum of squares, is calculated between V and V_R. The solution with the highest R^2 is then selected for each value of r. For choosing the minimum number of synergies required to represent the original matrix V, the curve of R^2 values versus synergies is fitted using a simple linear regression model, using all r synergies (see Figure 4 and Figure 5 for an example). The mean squared error[75] is then repeatedly calculated, each time removing the lower synergy point, until only two points are left or until the mean squared error falls below 10^{-5}. This strict constrain translates in choosing the most linear part of the curve and discarding it.

Originally, Lee and Seung applied NMF with image reconstruction purposes[56]. In this classical machine learning application, they could show that NMF was good for learning parts of objects. The procedure they employed was to convert black and white pictures to matrices containing levels of grey. To each pixel of a black and white picture, one can assign a value corresponding to how dark the pixel looks (e.g. 0=black, 1=white, all the included values being the levels of grey), as it is shown in panel D of Figure 2. By applying NMF, it is then possible to reconstruct the original photo with a quality which is gradually increasing with the number of synergies (or factorisation rank, describing the dimension of the basis vector). This is shown in Figure 3 and Figure 4. However, EMG signals are not different from a picture represented in levels of grey, if we think in terms of data representation.

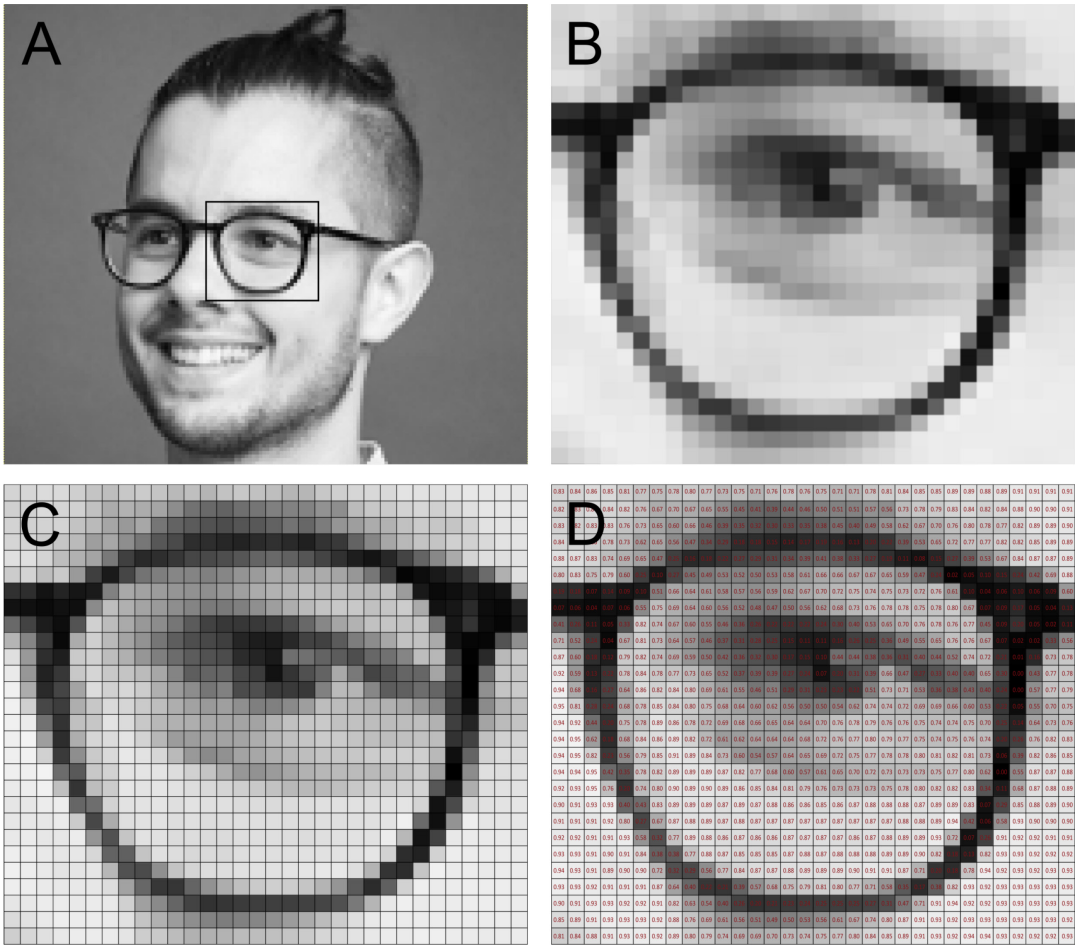

Figure 2 Graphical representation of a black and white picture as a matrix of pixels described by their level of grey (with black = 0, white = 1 and all the intermediate values representing the different levels of grey). Panel A: complete picture. Panel B and panel C: detail of the pixels considered as an example. Panel D: each pixel's level of grey is described by a number between 0 and 1.

Figure 3 The reconstruction quality, measured with the coefficient of determination R^2, increases when adding synergies to the model. It is of crucial importance to choose a method for determining the minimum number of synergies required to obtain a "satisfactory" reconstruction of the original data. One of these approaches is reported in the main text.

As mentioned above, many other NMF algorithms are available in literature and a lot is being done to constantly improve their computational performance both in terms of reconstruction capabilities and speed[75,78,89,81–88]. However, to date, the classical Gaussian approach is the most used for EMG decomposition[77,90,91] other than the most widely implemented in high-level programming environments.

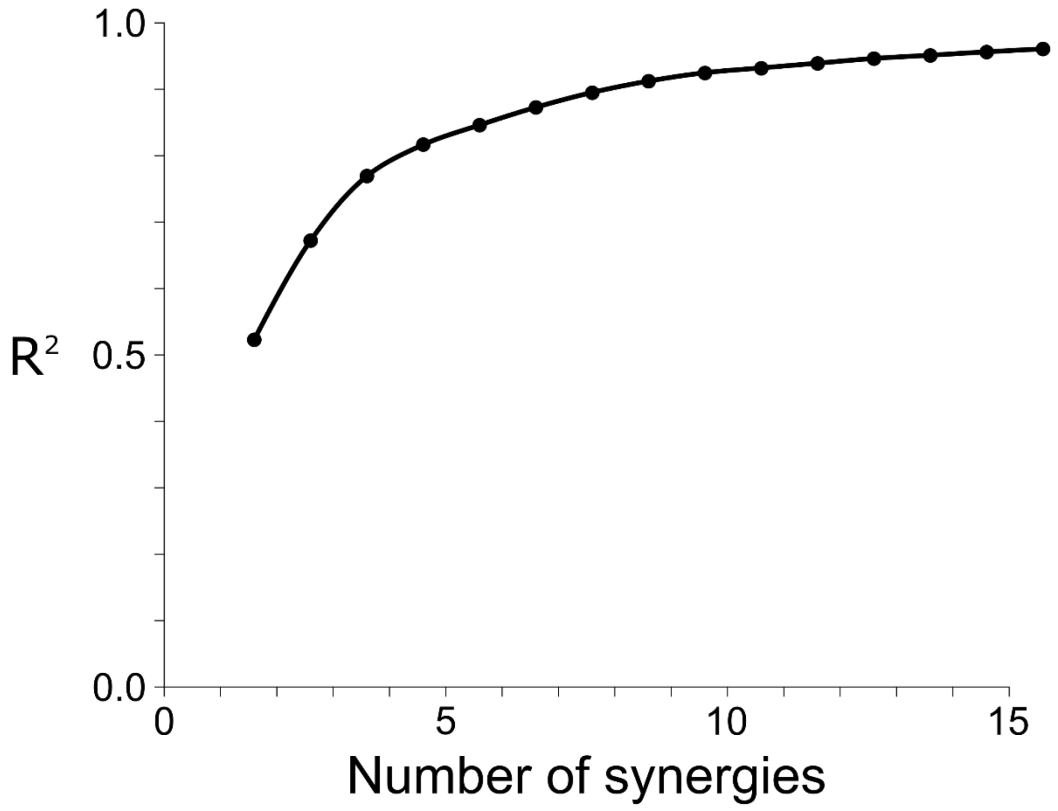

Figure 4 The reconstruction quality, measured with the coefficient of determination R^2, increases when adding synergies to the model. However, an asymptote is quickly reached and big changes in the number of synergies produce small changes in the reconstruction quality. The choice of the minimum number of synergies in the model is done by choosing a point of the curve above.

A typical factorisation of EMG activity appears as it is shown in Figure 5. The muscle activations matrix *V* is here graphically represented for seven muscles and a single gait cycle. However, it is common practice to concatenate several gait cycles in order to better account for step-to-step variability[92]. *V* can be approximated by the linear combination of *W* (the motor modules matrix) and *H* (the motor primitives matrix). Since it contains time-invariant coefficients, *W* is usually represented with bar graphs. On the contrary, *H* contains the evolution over time of the basic activation patterns and it is therefore better represented with time-dependent curves. When multiplying and summing synergy-by-synergy the elements of *W* and *H* following the concept reported in Equation 1, it is possible to reconstruct the original set of EMG data. The level of approximation, or reconstruction quality, is then defined as mentioned above, using the coefficient of determination R^2.

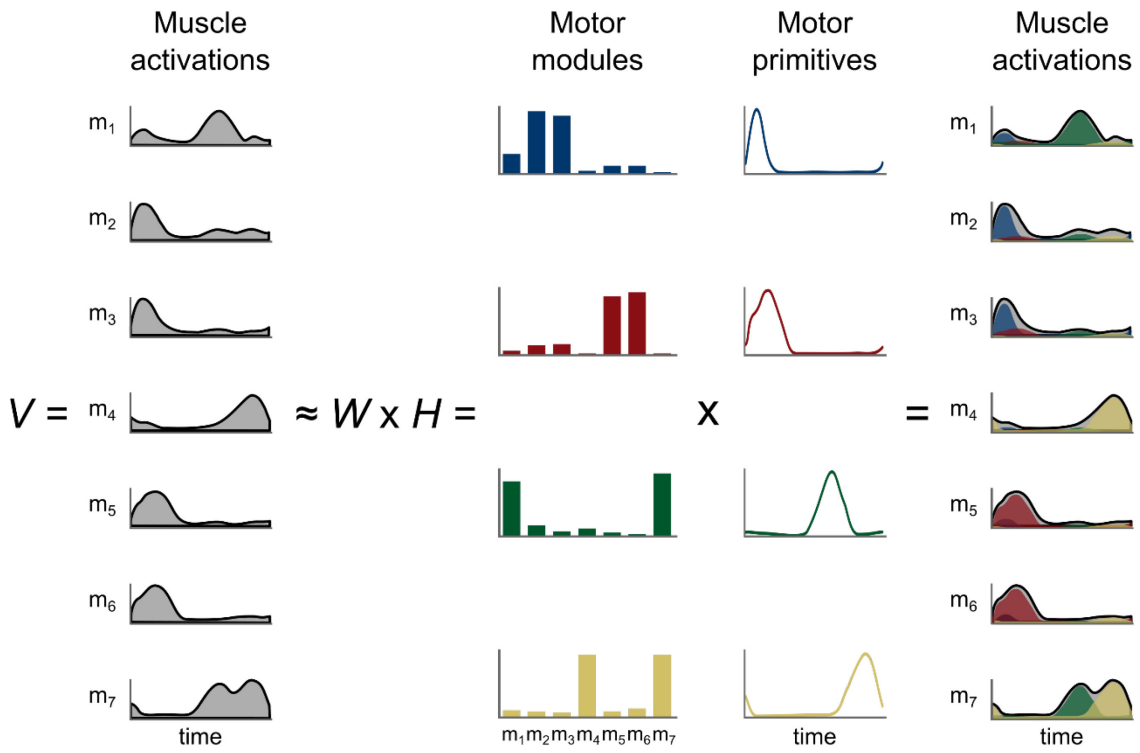

Figure 5 Graphical representation of EMG data factorised using the non-negative matrix factorisation approach. Seven muscle activations, indicated with m_n and part of the data matrix *V*, are compactly represented with only four synergies. The motor modules, indicated by the letter *W*, are the time-independent coefficients. The motor primitives, indicated by the letter *H*, are the time-dependent coefficients. The multiplication of *W* and *H* gives an approximate representation (reconstruction) of the original data *V*.

1.2.3 Muscle synergies for locomotion

In the past two decades, the scientific publications embracing the concept of muscle synergies have been flourishing and exponentially increasing in number. There are several examples of studies employing factorisation of EMG activity to study human locomotion.

For several reasons, the most widespread locomotion type that has been studied is walking[69,75,99–108,90,109–118,91,119–123,93–98]. Due to the easiness of examining this slow-speed type of locomotion, it is not a surprise that the majority of studies use walking as the main object of investigation. Also, it is clear that, contrarily to other locomotion types such as running, walking can be easily performed by patients, children and elderly and this feature notably extends the basin of potential participants. Nonetheless, running has been receiving increasing attention[69,75,77,91,100,104,124] as well. This might be partially due to the growing

popularity of distance running as a recreational sport activity over the last three decades[125]. Another reason to choose running over walking (or to study both conditions within the same experimental setup) is that, due to the different absolute and relative length of the stance and swing phases, different control mechanisms are likely to be used by the CNS[126,127]. Concerning this last matter, though, the field is still much open to new ideas, insights and exciting findings[126].

Unavoidably, the links between locomotion velocity and modular organisation have been investigated as well[91,96,97,100,103–105,112,122]. However, results are often contradictory and the reasons has not yet been clarified. Whether for computational or neurophysiological reasons, some studies found consistency in the recruitment of the same motor primitives and/or modules across varying velocities[91,96,97,112,122], while others found walking-, running- and/or velocity-specific sets of motor primitives and/or modules[91,100,103,105,112].

The role of muscle synergies for locomotion in pathology has been a focus of a few groups in recent years[72,94,123,128–132,98,99,101,103,107,112,114,119]. Given the simplification in presenting the data due to the dimensionality reduction, it is of course straight-forward to think to a possible clinical application of the method. There have been comparisons between healthy and Parkinson's disease[114,123,130], multiple sclerosis patients[99], spinal cord injury[72,98,131], cerebral palsy[94,107,129] and post-stroke[101,103,112,119,128] patients. However, as for the studies on the influence of velocity on the modular organisation of motion, also in pathology studies results are often difficult to interpret and require careful analysis.

The study of locomotion in perturbed conditions has as well started to meet the interest of many research groups[106,108,109,116,117,126]. The reasons for this choice are presented in the next paragraph, but it is straight-forward to grasp the importance of extending the controlled laboratory conditions to daily life.

Even if the consensus on factorisation techniques, data conditioning and interpretation is not unanimous, it is well accepted that human locomotion can be described with a small number of synergies. When analysing the EMG activities of lower limb's muscles[77,126], this number is usually equal to 4 or 5. A synergy might add when considering the upper body[75,91]. Figure

6 shows a typical factorisation output obtained from the EMG activity of lower limb's muscles during running.

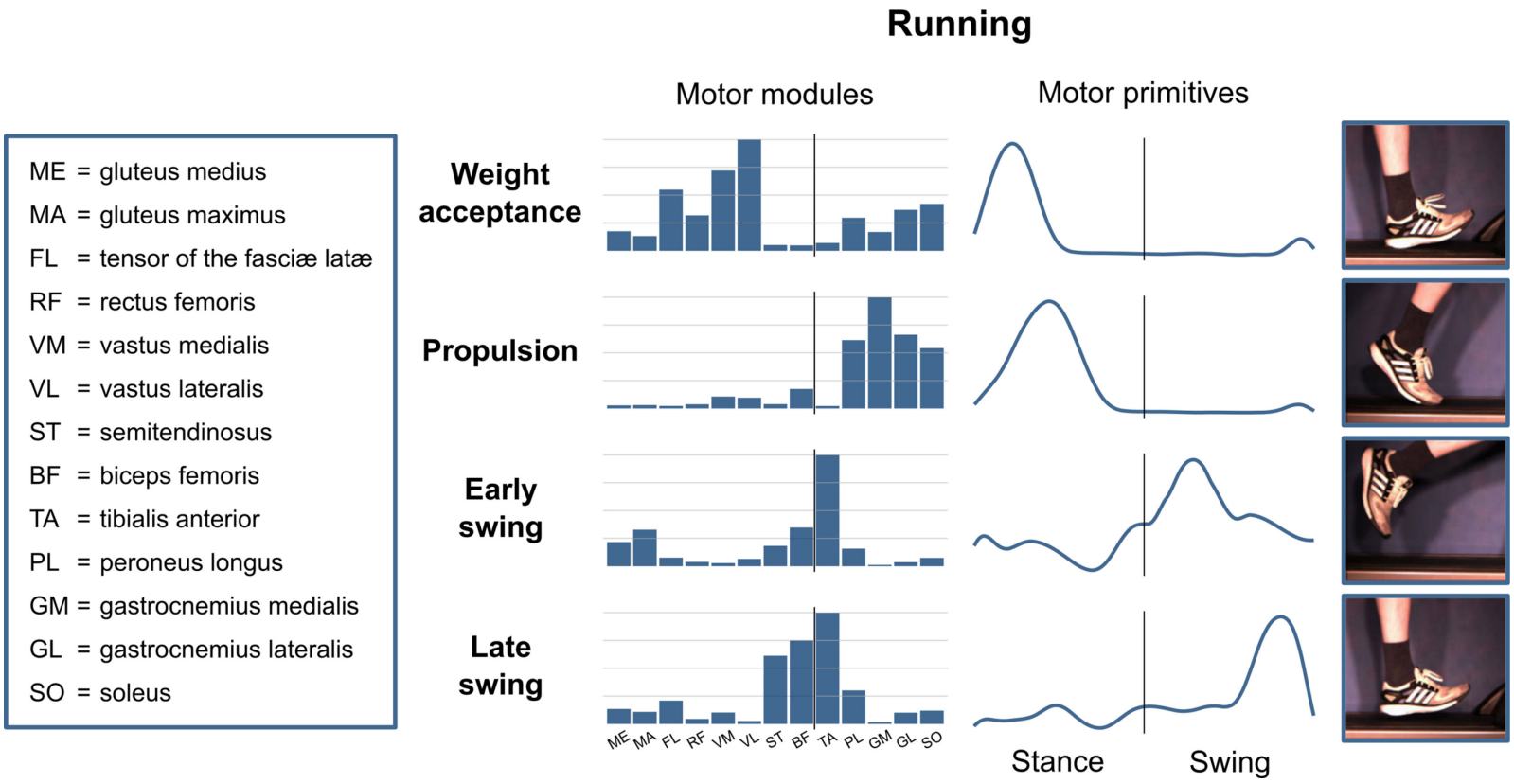

Figure 6 Typical EMG factorisation output for human locomotion. This specific example is related to treadmill running. Four muscle synergies describe as many phases of the gait cycle. The motor modules are presented on a normalised y-axis base. For the motor primitives, the x-axis full scale represents one gait cycle (stance and swing normalised to the same amount of points and divided by a vertical line) and the y-axis the normalised amplitude.

1.2.4 The link between muscle synergies and CPGs

As previously explained, muscle synergies are usually extracted from EMG data. This method, however, does not provide any information on where the movement and its control are generated. The neural connectivity underlying the cortical control of muscle synergies largely remains unknown[133,134]. Mapping and stimulation of cortical areas in the rhesus macaque have been used to obtain a representation of muscle synergies in the brain[134,135]. However, while synergies are easily identifiable in the motor cortex, they appear to not follow a rigorous (or decipherable) topographic organisation[134]. Similar perspectives arise when studying brain connectivity in humans[133]. Rana and colleagues could establish an interesting topography of pelvic floor muscle cortical control and its relation with muscle synergies[133]. Yet, these studies are all limited by the indeterminacy in attributing the structure of synergies to the various components of the CNS.

At the beginning of 2014, Levine and her colleagues published an important study aiming to link the modularity of task execution with the murine spinal interneurons[42]. They developed their ideas as follows. The organisation of a neural circuit can be outlined using viral tracing techniques based on retrograde transsynaptic transfer[136,137]. By infecting motoneurons with rabies virus, it is possible to obtain a tracing line which spreads from the infection site to only the infected cells, until the spinal cord[137]. Infecting the major muscles involved in joint movements (*medial* and *lateral gastrocnemii*, *tibialis anterior*, quadriceps, hamstrings, wrist flexors and extensors, elbow flexors and extensors), produces a spinal topography of the relevant interneurons. Levine and colleagues found this interneuronal pool to be mainly centred in the medial lamina V of the deep dorsal horn of the spinal cord[42]. They hypothesised that the stimulation of these premotor neurons would have led to "reliable and coordinated" motoneuron activity in the L2 and L5 spinal segments, providing a simple muscle synergy model[42]. By performing optical stimulation in the medial deep dorsal horn of premotor neurons specific to certain muscles, they could evoke coherent responses in the contextual spinal nerve roots[42]. On the contrary, the stimulation of non-specific interneurons could not produce specific responses, but only a mix of generalised activities[42]. This was the first experimental proof that some spinal neurons, which Levine and her co-authors named muscle "synergies encoders"[42], are responsible for encoding the motoneuron activation patterns for muscle synergies.

1.3 Locomotion as an ideal task for investigating movement

1.3.1 Treadmill locomotion

As already very well described by Bernstein in the late 1960s, locomotion is an ideal object for the investigation of movement[9]. Automatised, synergistic, general, cyclic and phylogenetically old, locomotion embodies many of the best boundary conditions a scientist could think of[9]. However, when it comes to the actual measurements, it happens very often that logistics, technology and time are not compatible with the study of free overground locomotion. Laboratories are usually too small to allow for recording of locomotion over long distances. EMG devices went wireless only recently and, when using optical motion capture, the researcher is always dependent on the acquisition volume, which is usually limited to a few cubic metres. For these and other reasons, treadmills are the most intuitive solution to overcome the natural obstacles presented by locomotion studies. But is it right to assume similarity between overground and treadmill gate?

Treadmill locomotion is often thought to be an invalid alternative to overground locomotion due to the mechanical advantage introduced by the moving belt. However, already in 1980, the Dutch biomechanist van Ingen Schenau showed that there are no mechanical differences between treadmill and overground locomotion as long as the belt's speed remains constant[138]. This "mystery" being unravelled, it is true that other factors might affect the patterns of treadmill locomotion: the compliance of the surface, the instability deriving from fixed rather than moving visual feedback, the degree of habituation, etc[139]. But in 1996, Jones and Doust demonstrated that the only causal factor responsible for differences between treadmill and overground running was the extra work required by air resistance rather than any mechanical variable[139]. Jones and Doust showed that to match the oxygen cost of overground running, a treadmill should be used with an inclination of 1% for a wide spectrum of steady-state running velocities (between 2.9 and 5.0 m/s) in athletes fully-habituated to treadmill running[139]. Subsequent studies found, however, runners[140] to be around 9% and walkers[141] around 23% more economical overground than on the treadmill, attributing the improvement to a lower ventilation and the excessive inclination of the treadmill (set at 1% as indicated by Jones and Doust). Other studies found differences in step length at fixed[142–145] and increasing[146] velocity, with shorter contact times and/or higher

cadence in treadmill compared to overground running. In general, however, there is the common opinion that the kinematic, kinetic and EMG differences between treadmill and overground locomotion are subtle enough to allow the use of treadmill for scientific purposes[141,143–145]. Also the muscle synergies seem to be little influenced when comparing treadmill and overground running. The differences have been proved to be mainly attributable to a shift in the timing (motor primitives) rather than a reorganisation of the muscle coefficients (motor modules)[147].

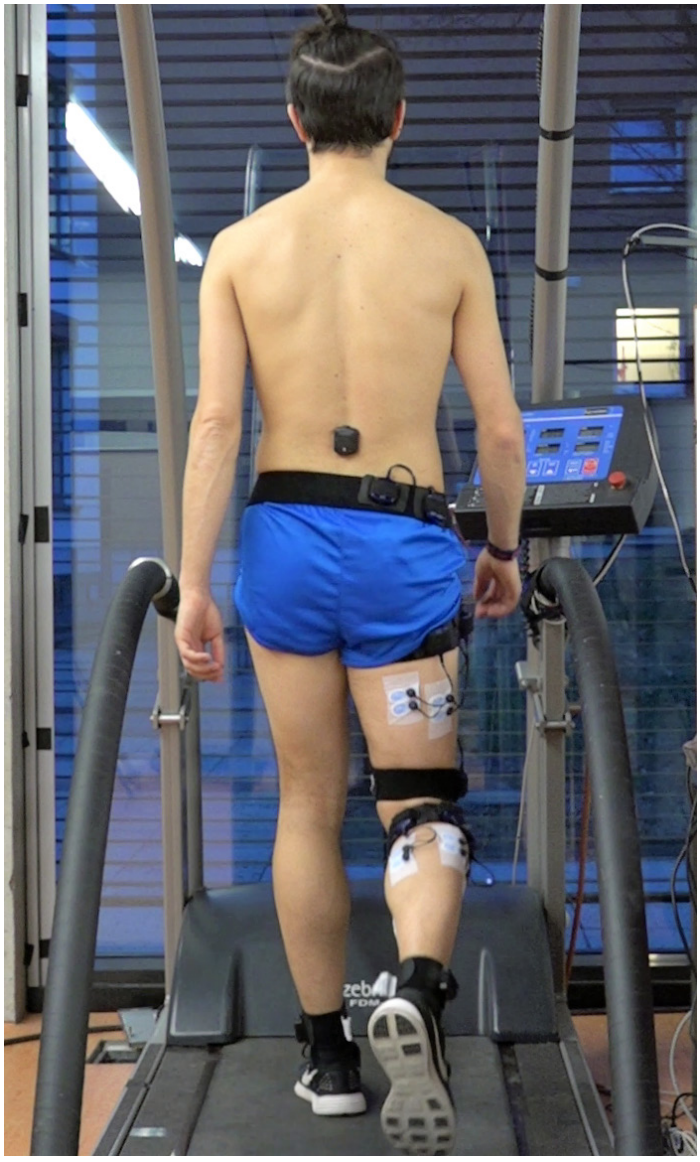

Figure 7 Typical treadmill experimental setup. The right leg is prepared for EMG acquisition and consecutive muscle synergies extraction. A pressure plate integrated in the treadmill is used to calculate some spatiotemporal gait parameters such as the foot strike index, contact and flight times, etc.

1.3.2 Perturbed locomotion

The mechanisms underlying the neural control of movement are not yet entirely understood[148]. Nevertheless, there is great amount of information on how animals move over even, solid surfaces such as treadmills or nice flooring of biomechanics laboratories[91,149–151]. Yet, daily life locomotion is far from resembling controlled laboratory conditions and implies facing an extraordinary amount of variables and interactions between them. All animals must regularly move in complex environments, source of constant external perturbations that make locomotion unavoidably unsteady. This unsteadiness causes a variety of behaviours such as changing speed, overcoming obstacles, dealing with slopes, recovering from external disturbances, negotiating variations in terrain, etc. Hunting and gathering, two daily activities that evolved in parallel with the genus *Homo*, inherently imply unsteady, long distance locomotion[152,153]. Being able to effectively move over diverse terrain conditions has been integral part of our evolution, not only to hunt and gather, but to escape predators and find mates as well[153].

One important way that animals use to optimise their locomotion is the minimisation of the energy cost[127,154,155]. Walking and running on irregular natural or artificial terrain as swamp, loam, stubble, grass, sand, snow, rubber, mountain trail and uneven-surface treadmill[156–165] has often been found to be more expensive than steady locomotion over even, solid surfaces. Yet, minimising the cost (or maximising the economy) of locomotion is not the only explanation for the changes observed in vertebrates' movement patterns when adapting to unsteady conditions[166]. Aiming to lower the energy cost of unsteady locomotor movements might not necessarily be the unique priority during locomotion. There might be, for instance, the need to perceive safety or comfort as parameters of similar or even higher importance than the optimisation of energy utilisation[166–169]. It has been shown that, during walking, humans undergo a decrease in their local dynamic stability and cope with perturbations by adjusting spatiotemporal gait parameters rather than the walking velocity[170–173]. There have been many attempts, both in humans and in animal models, to describe the biomechanics and energetics of balance control following single- and multiple-step perturbations[59,106,176–178,108,116,127,163,164,168,174,175]. However, the neuromuscular strategies employed by the CNS to cope with external perturbations are still poorly understood and this is where muscle synergies can strongly come into play.

Figure 8 The uneven-surface treadmill used in one of our studies (Woodway®, Weil am Rhein, Germany). The treadmill's belt consists of terrasensa® classic modules (Sensa® by Huebner, Kassel, Germany) aiming to simulate uneven ground conditions. At each step, an external perturbation is produced. Being the hills and valleys of the belt out of phase with respect to the gait, each perturbation is not predictable by the participant.

To overcome the non-predictability of uneven terrains, all animals have to constantly face changes in their locomotor patterns and must find quick solutions to maintain dynamic stability and functionality[168,179]. Since the beginning of last century, however, it has been clear that spinal CPGs are responsible for only the basic motor commands underlying locomotion and cannot generate that fine-tuned motor output required for facing environmental disturbances[2,12,15]. A core of excitatory (e.g. V2a), inhibitory (e.g. V2b and V1) and commissural (e.g. V0 and V3) spinal interneurons is involved in rhythm generation[7,45,46,180,181], left-right alternation[182] as well as flexor-extensor interaction[183,184]. Evidence that a finer, time-dependent tuning of the elementary CPGs commands might be of a supraspinal nature, has been found in the cat[185,186] and monkey[135]. Specifically,

neocortical and corticospinal tracts could influence the timing and level of locomotor activity[185]. There is as well indirect evidence, coming from observations in the monkey, that corticomotoneuronal cells in the primary motor cortex might be responsible for the creation of novel, highly skilled patterns of motor output in humans[187,188]. In the murine spinal cord, selective photo-stimulation of glutamatergic neurons is sufficient to initiate and maintain a locomotor-like activity[189]. When dealing with perturbations, though, the construction of movement must happen through a deep integration of all the available information, involving both spinal and supraspinal centres[79]. Thus, the role of proprioceptive afferent inputs must be of crucial importance for the control of perturbed locomotion. The two main functional classes of proprioceptive sensory receptors (or proprioceptors) are the group Ia/II muscle spindles and the group Ib GTOs[190].

Muscle spindles are receptors sensitive to muscle stretch. They detect changes in muscle's length and in the speed of these changes. Sensory information is canalised by primary type Ia and secondary type II sensory fibres around the muscle spindle, while motor action is taken up by gamma and beta motor neurons that activate the fibres within it. Any change in length and velocity due to stretch is detected by type Ia and II sensory fibres and transmitted monosinaptically to the alpha motoneurons of the involved muscle, which contracts to resist the stretch. The Ia afferent signal is contextually sent to Ia inhibitory interneurons that inhibit the alpha motoneurons of the antagonist muscles. If either the elongation or the velocity of it are excessive, the muscle spindles inhibit further stretching via reflex arcs.

GTOs are receptors sensitive to muscle tension. They are located at the origins and insertions of skeletal muscles and function as follows. When a muscle is activated, the relevant GTOs receive tension information that are sent to the spinal cord through type Ib sensory fibres. The information is relayed to the receptor-bearing muscle and is accordingly used to regulate the tension level. When tension exceeds certain boundaries, the so-called protective autogenic inhibition reflex, which is triggered by a form of negative feedback, inhibits further force production. This defensive mechanism results in a sudden relaxation of the involved muscle. However, during locomotion, the Ib input excites the interneurons rather than inhibiting them. Using positive feedback[191,192], this autogenic excitation regulates the generation of rhythm and, more specifically, the timing of the stance and swing phases[193].

In order examine the decoupling of spinal and supraspinal components for creating and controlling movement, Bizzi's group extracted the muscle synergies from the intact and spinal bullfrog's swimming and jumping[79]. They found most synergies to be shared, suggesting a central organisation[79]. However, an evident alteration of the time structure in the motor primitives let the authors conclude that sensory feedback might play a fundamental role in tuning muscle synergies[79]. It has been shown that a population of excitatory interneurons expressing the RORα orphan nuclear receptor are partially responsible for correction of foot movements in mice[194]. These corrective reflex movements are thought to be directed by an integration of touch information and descending motor commands from the motor cortex and cerebellum[194]. In support to this idea, the ablation of RORα interneurons in mice leads to an increase in foot slips during beam walking[194]. There is as well a population of inhibitory interneurons expressing the RORβ orphan nuclear receptor that has been proved to be responsible for fluid walking patterns in mice[195]. Abrogation of RORβ function results in a "duck gait" phenotype, characterised by exaggerated flexion patterns and strong degradation of the gait cycle general kinematics in an ataxic fashion[195]. Due to these observations, RORβ are thought to be regulating a low-threshold afferent inhibitory feedback circuit used to limit flexor motor activity[195]. Koch and colleagues proposed that RORβ interneurons "act as sensory filters to presynaptically gate proprioceptive afferent transmission and prevent abnormal flexor reflexes that disrupt the ongoing locomotor program, thereby securing the smooth rhythmic limb movements that are required for a fluid walking gait"[195]. Additionally, Akay and colleagues demonstrated that, upon elimination of proprioceptive feedback from muscle spindles in Egr3 (early growth response 3) knockout mice and of muscle spindles and GTOs in mutant (intersectional breeding of Pv::cre and Isl2::DTA) mice, walking and swimming patterns are drastically degraded[190]. In absence of proprioceptive feedback, the timing of ankle flexors during the swing phase is severely affected and directly impairs walking on a treadmill and on a horizontal ladder[190]. Also swimming patterns become extremely compromised, again in an ataxic fashion[190]. These important findings clearly show that the contribution of proprioceptive sensory feedback to the creation of coordinated movements is crucial as it is that of CPGs and, more generally, of the spinal and supraspinal components of muscle synergies. However, it is still unclear what exact role is played by sensory feedback in the generation and tuning of locomotor patterns[2,190].

1.4 Purpose of the thesis

This PhD project was divided into Work Packages (WP), tasks and milestones (◊). The main structure was composed by two methodological and two applicational studies. The first two aimed to establish some fundamental methods for analysing gait parameters by using an instrumented treadmill and the modular organisation of locomotion by using the concept of muscle synergies. With the second two studies the previously created methods for the study of unsteady locomotion conditions were implemented. Table 1 is a summary of the WPs, while Table 2 is a list of the milestones. Single tasks are omitted for brevity. A detailed description of the timeline and its contents is given below.

Table 1 Work Packages (WP) list. Each WP contained several tasks and sub-tasks, which are not reported here for brevity. The five WPs were divided in three main activity types (training, research and dissemination).

Work Packages	Activity description	Activity type
WP1	Training and development of generic and specific research skills.	Training
WP2	Experimental setups design, applying classical and new research methods.	Research
WP3	Reliability and accuracy evaluation of the experimental setups designed in WP2.	Research
WP4	WP2 and WP3 methodologies application in order to answer specific research questions.	Research
WP5	Dissemination of the gained knowledge and propagation of the scientific output to potential users.	Dissemination

Table 2 Milestones list. Each milestone indicates an important achievement in the PhD project.

Milestones	Title	Related WP(s)
◊1	Project kick-off.	WP1
◊2	WP2 complete.	WP2
◊3	WP3 complete.	WP3
◊4	WP4 complete.	WP3, WP4
◊5	PhD thesis dissertation.	WP5

1.4.1 Work Packages 1 and 2

The study of human locomotion is a fascinating way to appreciate the neurophysiological mechanisms underpinning the creation and control of movement. Walking is a daily activity for most of the population and running is quickly increasing its popularity among recreational sports[125]. Due to the highly-stereotyped movements on which they are built, walking and running certainly constitute an appealing investigation object for improving our understanding of the CNS's function[9]. However, the human body's motor control employed by the CNS for achieving and maintaining functionality, is not a trivial code to decrypt[1,9,12,52]. Amongst the few influential theories in motor control[66], the concept of muscle synergies, which is based on the idea of motor abundance, is one of the most intuitive and is living a profitable youth[52,53,65–72,57–64]. The fulfilment of the first two WPs was linked to the study and creation of the experimental and computational tools for: the assessment and analysis of spatiotemporal parameters of human gait; the measurement of EMG activities during human locomotion and the extraction of muscle synergies from EMG signals.

1.4.2 Work Package 3

The essential step before proceeding with the application of newly-created methods is a validation or, if this is not possible, a thorough examination of their reliability in terms of repeatability and reproducibility:

a) The first step consisted in writing an algorithm that, using plantar pressure distribution data, can automatically calculate the most interesting spatiotemporal parameters of treadmill walking and running (First study – A Pressure Plate-Based Method for the Automatic Assessment of Foot Strike Patterns During Running[196], page 30). This was an essential step also for correctly classifying and quantifying the foot strike patterns during different running conditions (e.g. see Third study – The Influence of Footwear on the Modular Organization of Running[77], page 72).
b) Then, two of the most used algorithms for the extraction of muscle synergies from EMG signals were implemented with the aim of assessing the influence on the output of the selective manipulation of data (i.e. filtering, intraday and interday reliability and factorisation approach; Second study – On the Methodological Implications of Extracting Muscle Synergies from Human Locomotion[75], page 48).

1.4.3 Work Packages 4 and 5

As mentioned in paragraph 1.3, the path towards a better understanding of human motor control should ideally pass through the analysis and comparison of steady and unsteady locomotion conditions[126,153]. Introducing perturbations in the system is a clever way to get familiar with the responses of a machine as complicated as the human body[126,153]. The fourth WP was designed according to this line of thought. Specifically, WP4 was divided into the two following main tasks:

c) Firstly, the effects of a small, acute perturbation on the modular organisation of running were analysed. Amongst the many possibilities to perturb fast locomotion, the choice fell on one of the most easily applicable modifications of running conditions: barefoot running. The recruited participants were all inexperienced barefoot runners. When they were asked to take their shoes off and run on an instrumented treadmill, they most likely underwent a perturbation given, amongst the other factors, by the lack of experience[77] (Third study – The Influence of Footwear on the Modular Organization of Running, page 72). Some upper body muscles were included in the analysis (total of 24 ipsilateral muscles), in addition to the standard lower limb setup used in the following study (which consists of 13 ipsilateral muscles).
d) Secondly, in order to increase the perturbation level and the randomness of it, the modular organisation of walking and running on an uneven-surface treadmill (see Figure 8) was assessed. The participants walked and ran on a normal treadmill and were then directed to the uneven-surface one. Both locomotion tasks were carried out without visual feedback (i.e. the participants were asked to not look at the foot placement), in order to ensure a proper perturbation[126] (Fourth study – Challenging human locomotion: stability and modular organisation in unsteady conditions, page 94). The standard lower limb's 13-muscle setup was used for the analysis.

The purpose of this thesis was to create, consolidate and apply those methods which are useful to provide novel insights into the neuromuscular control of perturbed human locomotion. Two automatic methods for gait parameters determination and extraction of muscle synergies from EMG signals were developed.

The hypotheses underlying the four studies presented here were the following:

i. An automatic method, based on plantar pressure distribution analysis, used to assess foot strike patterns during running is as reliable as the gold-standard technique in the field (i.e. video analysis).
ii. There is an influence of the EMG filtering conditions and NMF algorithm type on the outcomes of a muscle synergies extraction procedure.
iii. Barefoot running, considered as a form of perturbed locomotion, can induce changes in the modular organisation of movement.
iv. Locomotion over an uneven-surface treadmill can introduce a perturbation level which is high enough to trigger a response in the system which goes towards an increased robustness of the motor output.

The topics were addressed in four different published studies (WP5), which are presented in the body of this thesis. Each study is reported here with a reference style that matches the one of the whole thesis, in order to improve readability. References, figures, tables and equations are numbered consecutively for the same reason.

2 First study – A Pressure Plate-Based Method for the Automatic Assessment of Foot Strike Patterns During Running

Alessandro Santuz[1,2], Antonis Ekizos[1,2], Adamantios Arampatzis[1,2]

[1]Department of Training and Movement Sciences, Humboldt-Universität zu Berlin, Berlin, Germany
[2]Berlin School of Movement Science, Humboldt-Universität zu Berlin, Berlin, Germany

This is a post-peer-review, pre-copyedit version of an article published in Annals of Biomedical Engineering, 2016, 44 (5), 1646-1655. The final authenticated version is available online at: https://dx.doi.org/10.1007/s10439-015-1484-3.

2.1 Abstract

The foot strike patterns (FSP, description of how the foot touches the ground at impact) are recognized to be a predictor of both performance and injury risk. The objective of the current investigation was to validate an original foot strike patterns assessment technique based on the numerical analysis of foot pressure distribution. We analyzed the strike patterns during running of 145 healthy men and women (85 male, 60 female). The participants ran on a treadmill with integrated pressure plate at three different speeds: preferred (shod and barefoot, 2.8 ± 0.4 m/s), faster (shod, 3.5 ± 0.6 m/s) and slower (shod, 2.3 ± 0.3 m/s). A custom-designed algorithm allowed the automatic footprint recognition and FSP evaluation. Incomplete footprints were simultaneously identified and corrected from the software itself. The widely used technique of analyzing high-speed video recordings was checked for its reliability and has been used to validate the numerical technique. The automatic numerical approach showed a good conformity with the reference video-based technique (ICC = 0.93, p<0.01). The great improvement in data throughput and the increased completeness of results allow the use of this software as a powerful feedback tool in a simple experimental setup.

2.2 Introduction

For more than a century, scientists tried many ingenious ways to measure the distribution of pressure in the foot. The very first attempts with plaster-filled sacks[197] or spaded soil[198] at the turn of nineteenth century or the more sophisticated approaches using video recording[199] of the 1930s, had been as smart as difficult to apply and process. Since then, technology has made great strides leading, since the late 1970s, to a wide range of devices for the measurement of plantar pressure and ground reaction forces[200,201].

The popularity of distance running has greatly increased over the last three decades[125]. An average runner usually strikes the ground around three times per second[202]. The description of how the foot touches the ground during running, called foot strike pattern (FSP), depends on the location of the first contact area with the ground[203] and rearfoot (RS), midfoot (MS) and forefoot (FS) strike are the common classifications. If the FS is not followed by heel contact (like in the toe-heel-toe pattern), it is called toe strike (TS)[204]. FSPs during running have already been linked to injury risk[205–209] and performance[203,204,210–216]. Even if the conclusions are often contradictory and retrospective, it is clear that the study of FSPs is becoming increasingly connected not only to élite, but also to recreational sports.

The most used method to examine FSPs is the video analysis of the recorded landing[203,210,217–219]. This approach, however, is time-consuming and observer-dependent, since there is a need of trained people to manually scan the video recordings. Moreover, as it has been recently reported[219], the reliability of the observations decreases when the analyzed foot is not the one in the foreground. Another possible approach would be to make a kinematic analysis using a video system[220]. This solution can be a good asset when force plates are not available, but it certainly involves some preparation time to place the markers and calibrate the system. In addition, the compliance of the foot itself could produce some non-systematic errors[220]. Furthermore, the setup should include at least two cameras in order to include both feet in the analysis, thus adding some complexity to the experimental setup. These video-based methods, though, always rely on some level of manual preparation or elaboration and cannot be easily automated. Recently, also the use of inertial sensor has been proposed[221]. Given the inexpensive and lightweight nature of these devices, the approach is

certainly appealing for *in situ* applications. A validation with a kinematic method, though, showed a low reliability for TS cases[221]. Moreover, the strict requirements for sensors' supports stiffness and the quite complex post processing (i.e. filtering conditions), would require some experience and tuning before the proper application.

Other kinetic approaches, like the location of the center of pressure at impact[201], are widely used by pressure plate companies in their built-in software. This analysis alternative seems to be a good candidate for automating the evaluation process and the reason is twofold. First, it allows not only for a categorical classification of the FSP, but also for a quantification of it through the foot strike index (distance from the heel to the center of pressure at impact relative to total foot length). Second, it appears to be an appropriate metric when searching for correlation with injury risks[207,222]. Nevertheless, the analysis' boundary conditions must be clearly unfolded to the user, in order to: a) correctly interpret and report the outcomes and b) avoid inconsistencies like the presence of incomplete footprints when dealing with TS cases. The second point could be avoided by using pressure-sensitive insoles[223]. These devices, though, lack in reliability when dealing with lightweight participants and need a refined algorithm in order to detect the toe-off[223].

The existing methods are either time-consuming, observer-dependent or lacking in accuracy when trying to give a real-time feedback to the participant during treadmill running. Therefore, it is clear that a substantial standardization of the methods employed to evaluate the FSPs is currently missing.

The development of an automatic numerical algorithm able to classify accurately FSPs during running would provide several benefits: first, it would avoid any observer dependence by automating the evaluation process. This characteristic can considerably enrich the analysis, potentially including in a single evaluation step a big number of step cycles at once. Another important advantage is the ease of use: the researcher does not need to be trained to identify the FSP. Finally yet importantly, the automatic nature of this evaluation model can speed up the throughput of the outcomes, thus giving the chance to give online-feedback information about the foot strike.

The objective of the current investigation was to validate an original FSP assessment technique based on the numerical analysis of foot pressure distribution. Further, we aimed to make the method completely reproducible and therefore the boundary conditions and the calculation process are explained step-by-step.

We validated our custom algorithm against one of the most used and accepted methods, the video analysis[203,205,210,215,219,224], using a large sample size. The investigation was conducted across a wide spectrum of submaximal speeds and two different conditions (barefoot and shod) in order to prevent the analysis of only very specific conditions. Finally, the treadmill approach ensures a consistent number of gait cycles to analyze and average. This significantly improves the impact of the method compared to others considering only one or two steps by using fixed cameras while running over ground or force plates[203,204,226,227,210,216–220,224,225].

2.3 Materials and methods

2.3.1 Experimental design

We recruited people of both genders and various ages (n = 145; 85 male, 60 female, see Table 3 for details). The criteria for assessing their running experience were defined as follows. Inexperienced (n = 57; 29 male, 28 female): people that were inactive, doing other sports[228] or that just took up running[219] (for a period ≤ 1 year). Recreational (n = 67; 41 male, 26 female): people running more than 20 km/week[204,229] and averaging three or more sessions/week in the past 5 years[230]. Competitive (n = 21; 15 male, 6 female): athletes running more than 40 km/week[213,228], registered in athletics or triathlon clubs and competing at the regional, national or international level in any event, except throwing and walking[231].

All of them gave informed consent to the experimental procedure, according to the rules of the local scientific board. None of the participants showed or reported any history of neuromuscular or musculoskeletal impairments at the time of the measurements. Moreover, in the six months prior to the measurements day, none of them has suffered any injury to the lower limbs and they all reported to be habitually shod during daily life and, if applicable, when running. No participant was an experienced barefoot runner.

Table 3 Participants' anthropometric characteristics and running velocities. Values reported as mean ± s.d.

			M	F
n	**Total**	145	85 (59%)	60 (41%)
	Inexperienced	57	29	28
	Recreational	67	41	26
	Competitive	21	15	6
Height [cm]	**Total**	175 ± 9	180 ± 7	168 ± 6
	Inexperienced	173 ± 9	178 ± 8	168 ± 6
	Recreational	176 ± 9	181 ± 6	169 ± 7
	Competitive	176 ± 7	180 ± 5	168 ± 5
Body mass [kg]	**Total**	69 ± 11	74 ± 9	62 ± 8
	Inexperienced	68 ± 11	74 ± 11	62 ± 8
	Recreational	70 ± 11	76 ± 9	62 ± 8
	Competitive	67 ± 9	71 ± 6	57 ± 5
BMI [kg/m^2]	**Total**	22 ± 2	23 ± 2	22 ± 3
	Inexperienced	23 ± 3	23 ± 3	22 ± 3
	Recreational	23 ± 2	23 ± 2	22 ± 3
	Competitive	21 ± 1	22 ± 1	20 ± 1
Age [years]	**Total**	30 ± 9	32 ± 9	27 ± 8
	Inexperienced	29 ± 10	30 ± 10	27 ± 10
	Recreational	32 ± 8	34 ± 8	29 ± 8
	Competitive	27 ± 6	28 ± 6	25 ± 5
Speed pref. [m/s]	**Total**	2.8 ± 0.4	3.0 ± 0.5	2.6 ± 0.3
	Inexperienced	2.6 ± 0.3	2.7 ± 0.3	2.5 ± 0.2
	Recreational	2.8 ± 0.3	3.0 ± 0.3	2.6 ± 0.2
	Competitive	3.5 ± 0.5	3.7 ± 0.4	3.0 ± 0.1
Speed faster [m/s]	**Total**	3.5 ± 0.6	3.7 ± 0.6	3.1 ± 0.3
	Inexperienced	3.1 ± 0.4	3.3 ± 0.4	3.0 ± 0.3
	Recreational	3.5 ± 0.4	3.7 ± 0.4	3.2 ± 0.2
	Competitive	4.4 ± 0.6	4.7 ± 0.3	3.6 ± 0.1
Speed slower [m/s]	**Total**	2.3 ± 0.3	2.4 ± 0.4	2.1 ± 0.2
	Inexperienced	2.1 ± 0.3	2.2 ± 0.3	2.0 ± 0.2
	Recreational	2.2 ± 0.3	2.4 ± 0.3	2.0 ± 0.2
	Competitive	2.7 ± 0.3	2.8 ± 0.3	2.4 ± 0.1

The FSP assessment method was double: a video analysis using a high-speed camera and an automated evaluation using an algorithm to interpret pressure distribution data. After checking its reliability by comparing the outcomes of eight different observers, the video analysis was considered as a reference technique for validating the custom software.

2.3.2 Material

The foot strike videos were recorded using a high-speed video camera (Flare 4M180-CCL, IO Industries Inc., London, Canada) operating at 550 Hz, with dedicated recording software (Simi Grab 2.1.1, Simi Reality Motion Systems GmbH, Unterschleissheim, Germany). The resolution was set to 984x400 pixels. Pressure distribution patterns were recorded at 120 Hz through a pressure plate (FDM-THM-S, zebris Medical GmbH, Isny im Allgäu, Germany) integrated in a treadmill (mercury, H-p-cosmos Sports & Medical GmbH, Nussdorf, Germany). The pressure plate data were acquired using the proprietary software (WinFDM-T v2.5.1, zebris Medical GmbH, Isny im Allgäu, Germany) and then extracted in a raw format for autonomous post-processing (R version 3.1.2, R Foundation for Statistical Computing, R Core Team, Vienna, Austria). Both the camera and the plate were synchronized using an analogue signal. The camera was set up 350 cm laterally to the left side of the treadmill, mounted on a tripod at a height of 29.5 cm and angled perpendicular to the sagittal plane of the subjects.

2.3.3 Protocol

On the treadmill, the participants conducted a self-selected warm-up[216,232], in order to choose a comfortable running pace. The procedure to find the comfortable speed was implemented using the method of limits[233]. The speed was randomly increased with steps of 0.02 to 0.05 m/s at varying time intervals (around 5 to 10 s) until the participant was comfortable with a specific pace. Then the operation was repeated starting from a faster speed (around 0.5 to 1 m/s higher) and randomly decreasing it as previously done. If the comfortable value was not differing more than 10% from the previous, the average of the two values was taken as the preferred. Otherwise, the whole procedure was reiterated. The warm-up protocol typically lasted between 5 and 10 minutes, with an average of 6.8 ± 1.2 minutes.

After being instructed about the protocol, the participants ran for 90 s in four different conditions: preferred speed (shod, 2.8 ± 0.4 m/s), faster speed (shod, 3.5 ± 0.6 m/s) slower speed (shod, 2.3 ± 0.3 m/s), preferred speed (barefoot). For competitive and recreational runners, the faster speed was then determined calculating the average speed maintained

during the personal best time effort over a 10 km race (mean 125 ± 7% of preferred speed). If no information was available (e.g. for inactive participants), the faster speed was set as the 125% of the preferred. The slower speed was set as the 75 to 85% of the preferred speed (mean 80 ± 4%), accordingly to each participant's preference. As a guideline to identify the slower speed, the participants were asked to report a maximum value of 2 (i.e. weak, light effort) on the modified Borg Rating of Perceived Exertion scale[234]. Each condition was repeated, thus giving eight datasets for each subject. The order of the eight trials was completely randomized. Before every trial, each participant performed 60 s of familiarization running on the treadmill, in order to allow for accommodation[235]. Following, a dataset of 30 s was recorded. There were no stops between the trials, except the one or two necessary brakes to take the shoes off before running in the barefoot condition.

2.3.4 Analysis

Two main parts formed the FSP analysis: the video analysis as a well-accepted reference technique and the pressure plate data elaboration through a new numerical computation algorithm.

For the video analysis, eight observers were trained to identify the three FSPs (i.e. RS, MS and FS) by showing them ten trivial (clearly identifiable) and ten non-trivial videos taken outside the present study. A trivial video is namely a representation of an unambiguous FSP (e.g. a strong RS, where the rearfoot unequivocally touches the ground before any other part of the foot). On the contrary, a non-trivial video is a recording where the FSP is not clearly identifiable (e.g. a MS or a light FS or RS, where the gap between the outsole and the ground is not evenly assessed among different observers because of lighting conditions, superposition of similar colors, etc.). After becoming able to classify all the subjected data, the observers were asked to look at the study's videos and to recognize the RS, MS and FS cases. As reported from other studies[219], there is indeed a decline in reliability of around 10% when making a video analysis of the foot in the background. To avoid this potential source of measurement uncertainty, only the data related to the left foot (foreground image) are presented.

A typical analysis scenario is shown in Figure 9. We report here only the data related to observations of the foreground foot (left), in order to avoid any additional measurement uncertainty[219] due to the difficult interpretation of a background image.

The pressure plate-data elaboration revolves around the concept of strike index (SI). The SI, as originally defined by Cavanagh and Lafortune[201] and then adopted from several other authors[204,214,220,221,225,236–240], is the distance from the heel to the center of pressure at impact relative to total foot length. However, the most important assumption made from this method is to know the foot (or shoe) length. The plate itself cannot measure this quantity accurately, since in cases like the TS the entire foot does not touch the treadmill. Therefore, for a clear identification of the FSPs from the pressure distribution data, the first step is to determine for each participant the footprint lengths. The measured lengths with and without shoes (for shod and barefoot trials, respectively) have been used as a reference to carry on the analyses. A custom-made caliper has been used for the measurements. The bare foot length has been considered as the distance from the pternion point to the most anterior point of the longest toe, measured parallel to the foot axis. In a similar way, the shoe length was measured as the distance between the perpendicular projection to the ground of the most posterior and the most anterior points of the outsole. These values constitute the foot and shoe measured lengths. The information regarding the footprint length is necessary especially when dealing with incomplete footprints during running (like in the case of TS, where the heel never touches the plate), since the pressure plate does not give any information about how long the original footprint should be.

To identify the forefoot, the midfoot and the rearfoot, the footprint is divided into three geometrically equal parts, each representing one third of the total length. Using the pressure values of the individual foot recorded from the plate, the code evaluates the footprint length (calculated length) along the treadmill's anterior-posterior axis. If the calculated value differs more than 5% from the measured one (e.g. for TS cases), the footprint is corrected with the "real" value, like shown in Figure 10. The most important assumption underlying this step is that, during the toe-off phase, the tip of the shoe or of the foot always touches the ground. The width of the foot is considered as the widest footprint recorded. The footprint is thus localized within its real boundaries: this is done by expanding each pressure matrix in length and width with the appropriate number of zero elements in order to reach the correct length and width.

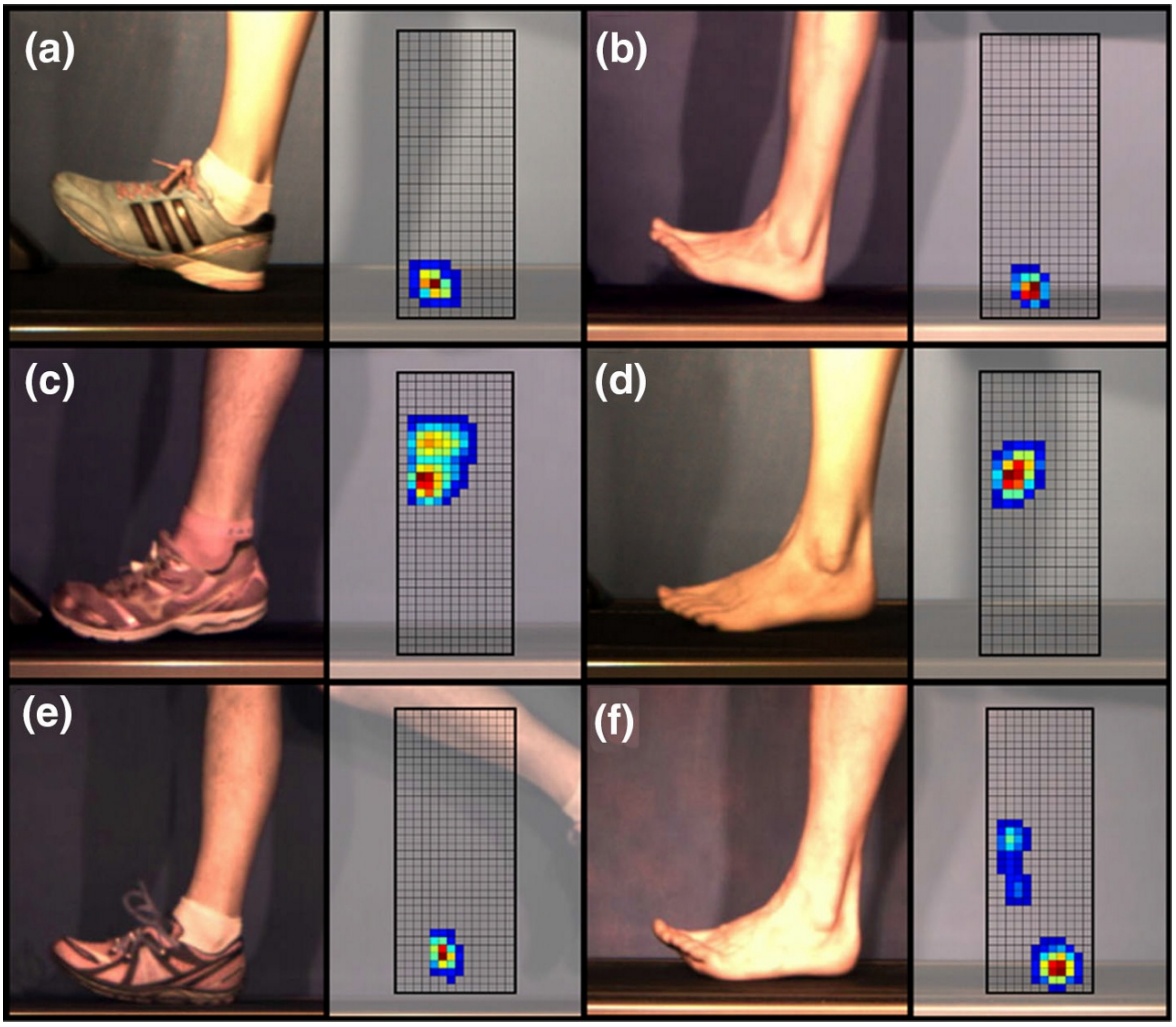

Figure 9 A typical video analysis scenario and the correspondent pressure distributions at impact. Pictures (a) and (b) represent the rearfoot strike, while (c) and (d) show the forefoot strike. Pictures (e) and (f) show the difficulty of assessing the FSP using the video analysis.

The calculation of the SI, then, automatically provides one of the three FSPs (RS, MS or FS, being the TS case included in the FS). To temporally locate the impact, the first recorded data after the swing phase has been taken as a reference, thus defining "impact" as "initial contact"[201]. In the algorithm this is considered as the first nonzero pressure matrix after the last toe-off. In Figure 11, a flowchart shows the logic of the FSP determination algorithm, emphasizing the individual steps and their interconnections.

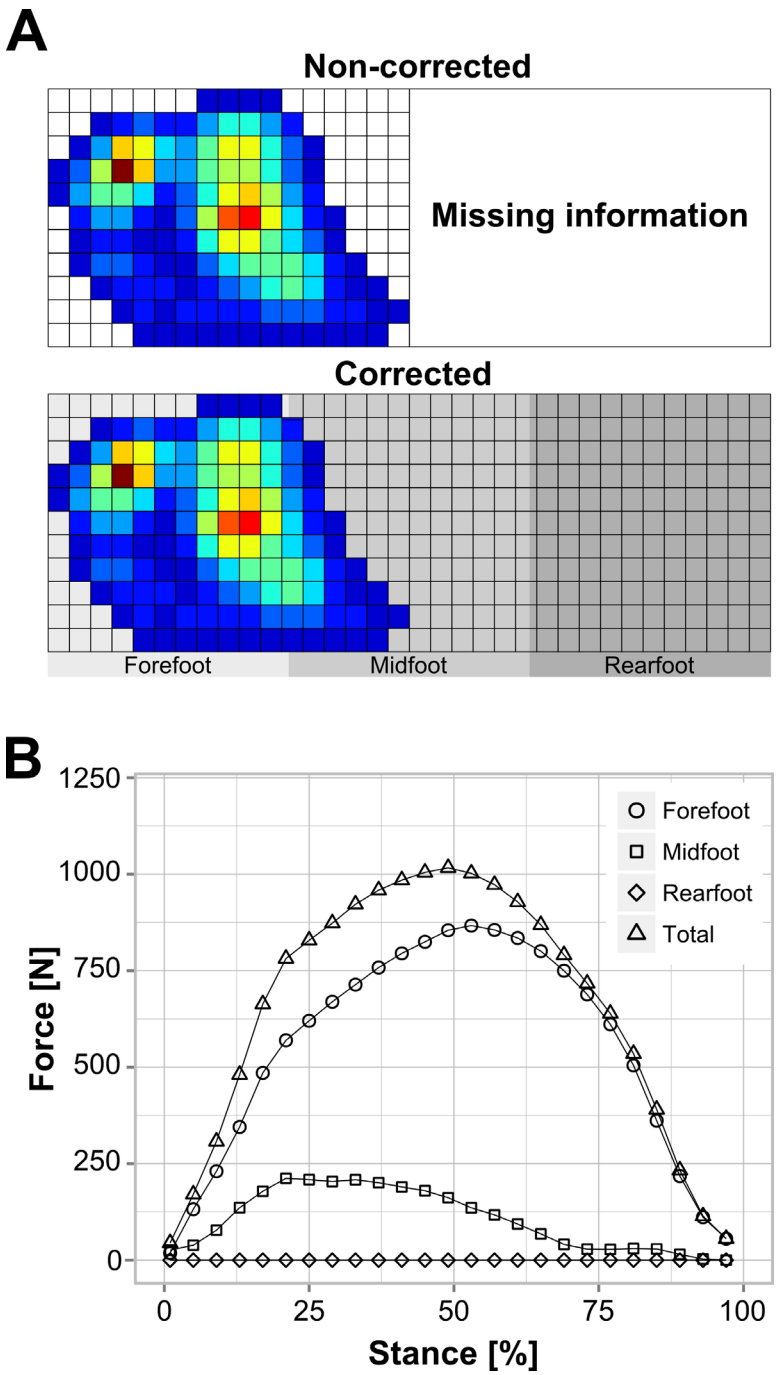

Figure 10 Graphical representation of the footprint's length correction. In this strong toe strike case, the participant never touched the ground with the rearfoot, as shown in the pressure distribution (A) and in the ground reaction forces (B) graphs. The identification of fore-, mid- and rearfoot is possible only after the footprint correction via software (A).

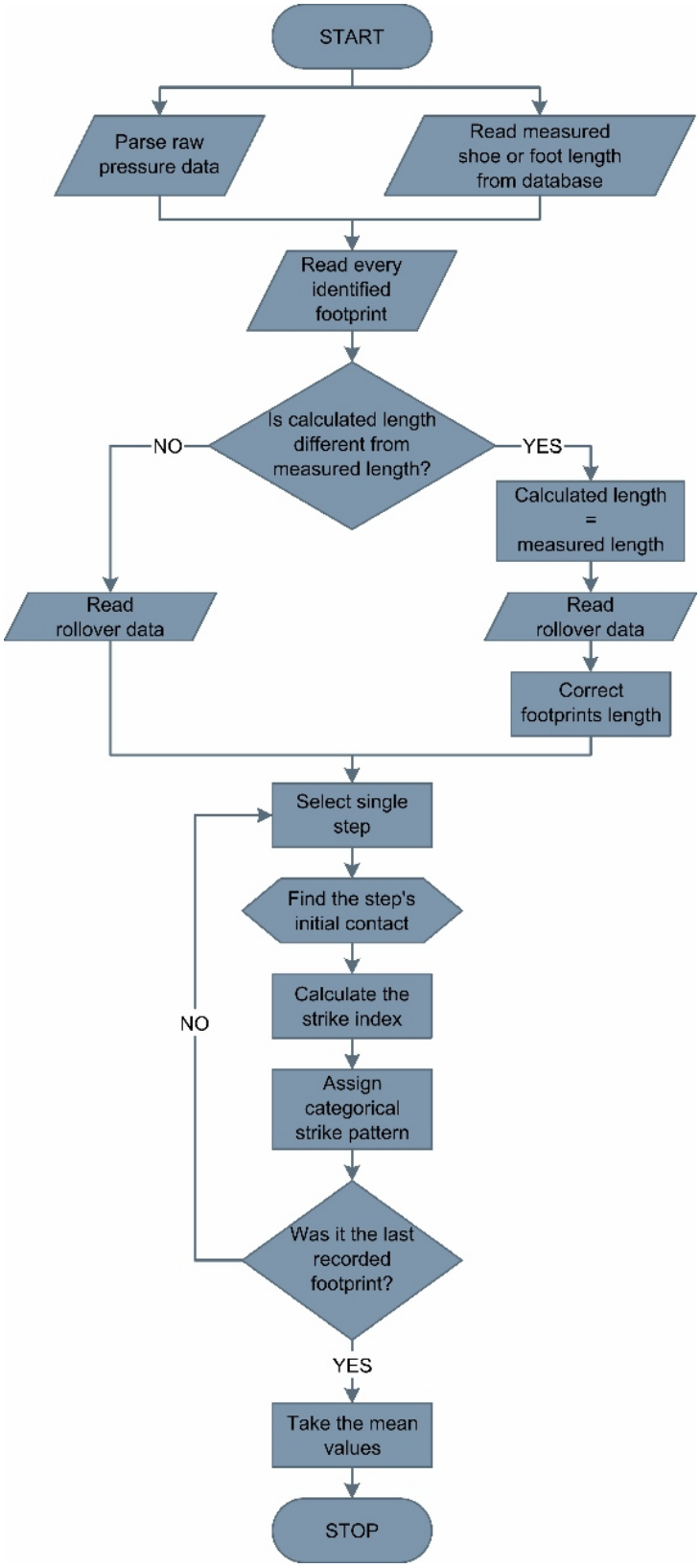

Figure 11 Flow chart showing the logic of the FSP determination algorithm. Every fundamental step and interconnection are reported for reproducibility purposes.

2.3.5 Statistics

We calculated a two-way Intraclass Correlation Coefficient (ICC) for single measurements to assess the agreement between eight different observers conducting the video analysis.

The 95% confidence interval of our sample's margin of error was estimated through a bootstrapping procedure, in order to have an indication about the uncertainty of our FSP estimate. The original data set was resampled 10000 times with replacement, considering the FSP as the main parameter. Starting from a sample size of 41 in order to consider the sampling distribution to be normal (central limit theorem), the procedure was repeated until the total sample size was reached, with increasing steps of 10. The minimum sample size was chosen since the central limit theorem states that the sampling distribution of the statistic is normal for sample sizes greater than 40. After elaborating the data, the agreement between the video analysis and the numerical approach was calculated, thus comparing our algorithm with a reference technique.

2.4 Results

As shown in Table 4, the agreement between the eight observers is usually higher for RS cases in all conditions (ICC values from 0.83 to 0.96, confidence level 0.99). The FS pattern assessment suffers a decrease in ICC value for the shod condition at faster and slower speeds (ICC values 0.65 and 0.73, respectively), while it produces higher agreement in the other conditions (ICC values 0.86 for the shod condition, preferred speed and 0.92 for the barefoot case). In evaluating MS cases, the observers never reach high agreement (ICC values 0.51 to 0.64). When joining the midfoot and forefoot strikes into a single pattern (MFS), thus identifying only two types of FSP, the ICC inevitably increases for all conditions (0.83 to 0.96). For this reason, the video analysis can be a proper reference method only when considering the RS and the joint MFS patterns.

Table 4 Intraclass correlation coefficient (ICC) calculated between the outcomes of eight observers' video analyses.

Condition	Speed	FSP	ICC	Lower bound	Upper bound	*p*-value
Shod	Preferred	RS	0.83	0.77	0.87	<0.01
		MS	0.58	0.50	0.67	<0.01
		FS	0.86	0.82	0.90	<0.01
		MFS	0.83	0.77	0.87	<0.01
	Faster	RS	0.86	0.82	0.90	<0.01
		MS	0.51	0.42	0.60	<0.01
		FS	0.65	0.57	0.72	<0.01
		MFS	0.86	0.82	0.90	<0.01
	Slower	RS	0.89	0.86	0.92	<0.01
		MS	0.64	0.56	0.72	<0.01
		FS	0.73	0.66	0.79	<0.01
		MFS	0.89	0.86	0.92	<0.01
Barefoot	Preferred	RS	0.96	0.94	0.97	<0.01
		MS	0.53	0.44	0.62	<0.01
		FS	0.92	0.89	0.94	<0.01
		MFS	0.96	0.94	0.97	<0.01

As reported in Table 5, the video analysis, averaged among eight observers, found 76.7% RS cases. The joint MFS pattern, accordingly, constitutes the 23.3% of the total 1160 observations (145 participants, five conditions, two trials for each condition). Being the measurement uncertainty of the numerical analysis independent on the foot considered, the results are presented for both the left and the right feet.

Combining all the results (preferred, faster and slower speed shod and preferred speed barefoot), the numerical analysis found 78.3% left and 77.8% right RS cases (see Table 5). Joining again the MS and FS patterns into a single one (MFS), makes the new pattern to contribute for 21.6% (left) and 22.2% (right) of the total observations. The computation times are 0.45 s for every second of recorded data on Intel® Core™ i5-5250U @ 2.70 GHz with 8 GB RAM on Windows 7 64-bit and 0.15 s on Intel® Xeon® X5650 @ 2.66 GHz with 48 GB RAM on Windows 7 64-bit. To assess the validity of the numerical approach, the outcomes of the RS and MFS patterns were compared with the video analysis. The investigation of RS and MFS has been chosen because in these two FSP cases the reliability of the video analysis was very high. The ICCs, related to the left foot results, are calculated on the number of observations for each FSP. The RS and MFS cases return high values (0.93). All agreements are significant ($p<0.01$).

Table 5 Comparison of the video analysis' outcomes with the numerical results. FSP = foot strike pattern, RS = rearfoot strike, MFS = midfoot and forefoot strike joint). The video analysis' results are an average of the eight observers' outcomes over all data (preferred, faster and slower speed shod and preferred speed barefoot). In the last two columns, the ICCs (and the relative p-values) between the video and numerical results are reported.

FSP	VIDEO	NUMERICAL		ICC (Video vs. numerical)	Lower bound	Upper bound	*p*-value
	Left	Left	Right				
RS	890 (76.7%)	908 (78.3%)	903 (77.8%)	0.93	0.91	0.94	<0.01
MFS	270 (23.3%)	252 (21.6%)	257 (22.2%)	0.93	0.91	0.94	<0.01

The 95% confidence interval estimation of our dataset's standard errors gave the values reported in Figure 12. The margin of error was calculated in order to estimate the likelihood of obtaining results close to the whole population's. The following lines specify the values used to create the histogram in Figure 12. The four conditions' video analysis presented the following FSPs: shod, preferred speed, 88.3% RS and 11.7% MFS; shod, faster speed, 87.6% RS and 12.4% MFS; shod, slower speed, 84.8% RS and 15.2% MFS; barefoot, preferred speed, 46.2% RS and 53.8% MFS. The numerical analysis of the left footprints produced: shod, preferred speed, 89.0% RS and 11.0% MFS; shod, faster speed, 89.0% RS and 11.0% MFS; shod, slower speed, 86.9% RS and 13.1% MFS; barefoot, preferred speed, 49.0% RS and 51.0% MFS. The numerical analysis of the right footprints produced: shod, preferred speed, 88.3% RS and 11.7% MFS; shod, faster speed, 87.6% RS and 12.4% MFS; shod, slower speed, 86.2% RS and 13.8% MFS; barefoot, preferred speed, 49.0% RS and 51.0% MFS. The related margins of errors were 5.2%, 5.4%, 5.8% and 8.1% for the video analysis; 5.0%, 5.0%, 5.5% and 8.1% for the numerical analysis of the left footprints; 5.0%, 5.1%, 5.4% and 8.1% for the numerical analysis of the right footprints. The relative standard errors values are 2.6%, 2.7%, 3.0% and 4.1% (video analysis); 2.6%, 2.6%, 2.8% and 4.2% (numerical left); 2.6%, 2.6%, 2.8% and 4.2% (numerical right).

Table 6 shows the FSP distribution during shod and barefoot running at a comfortable speed (here called "preferred speed").

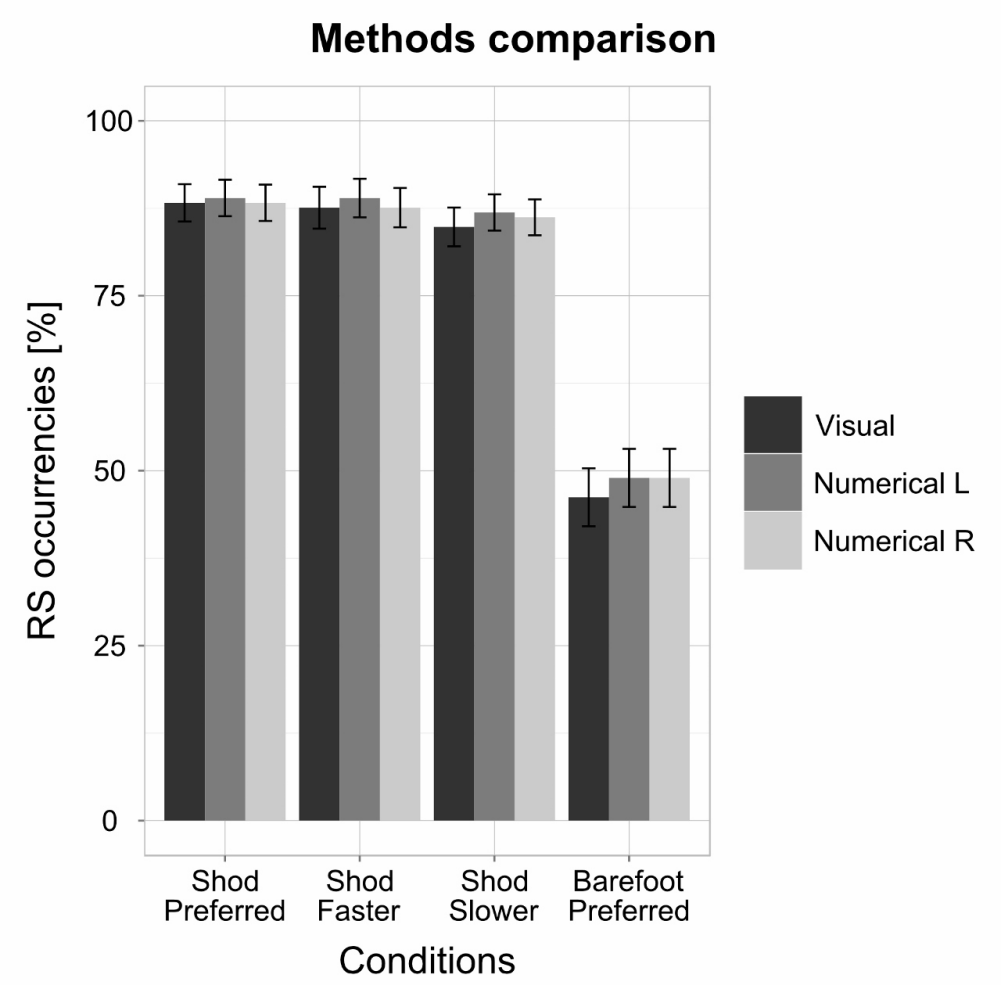

Figure 12 Percentage of rearfoot strikes (RS) detected with the video analysis (averaged among eight observers) and by the numerical method (for both left and right foot). The error bars represent ± se (standard error) values, calculated through a bootstrapping procedure.

Table 6 Foot strike patterns occurrences during shod and barefoot running at preferred speed. The results are obtained using the objective numerical algorithm. FSP = foot strike pattern, RS = rearfoot strike, MS = midfoot strike, FS = forefoot strike.

FSP	Shod		Barefoot	
	Left	Right	Left	Right
RS	89.0%	88.3%	49.0%	49.0%
MS	9.6%	10.3%	42.8%	44.1%
FS	1.4%	1.4%	8.2%	6.9%

2.5 Discussion

This study aimed to create and validate an automatic method to evaluate foot strike patterns during running. Using the plantar pressure distribution recorded by a pressure plate integrated in a treadmill, we created an algorithm able to detect the FSP during running. This provides the basis for the development of an online-feedback system. To validate the method, its agreement with a solid and often used reference technique - namely the analysis of the sagittal plane video of the striking foot - was checked. Among eight independent observers, the reference technique provided a very reliable RS and MFS patterns recognition. Therefore, it was considered an adequate gold standard for the evaluation of the new algorithm. The algorithm is able to quickly and accurately recognize the FSP and can therefore be easily used as a fast feedback tool. On a standard machine, the computation time is 45% of the recorded time (see results for details).

FSP assessment through video analysis showed a very high reliability in identifying RS cases. The rearfoot is most of the times easily located by the observers and the common strong dorsiflexion associated with a RS pattern[204] helps in clearly defining the case. For similar reasons, also the FS case is often clearly isolated, especially when the plantar flexion right before the strike is substantial. There is, however, a low conformity among observers in determining MS cases. This is because the midfoot is often difficult to locate and often leads to the misinterpretation of the FSP. In addition, it is not always trivial to identify a gap or a contact between the outsole (or the bare foot) and the treadmill. These factors add an amount of uncertainty that translates in a lowered agreement between observers (low ICC). Joining, as we propose, the MS and FS cases into one single pattern called MFS solves most disagreeing cases. The decrease in ICC values for the shod condition at faster and slower speeds also suggests a dependence of the inter-rater reliability on the running speed. Since the foreground foot was the left, the image of the right foot was often difficult to interpret. A previously-reported decline in reliability when analyzing the foot in the background[219], convinced us to only present the data related to the image in the foreground, namely the left foot. Additionally, the video analysis is not adequate for giving online-feedback information about the FSP and proved to be extremely time-consuming, especially when the number of trials is not small. The video analysis, one of the most popular methods for FSP

assessment[203,205,210,215,219,224], is anyhow solid enough to be considered as a reference technique for validating our algorithm. The uncommonly big sample size of 145 participants was chosen in order to reduce the chance of biases in our measurements. In these kinds of studies, it is very common to use small sample sizes[215,216,220,221,224,225,227,238]. This is partly due to the complex structure of some experimental setups. The chosen sample size can be evaluated by estimating the margin of error. This quantity contains the information about the uncertainty with which one predicts to describe the whole population. This means that our sample estimate will not differ from the true population's by more than the margin of error values 95% of the time (the chosen confidence level). This is obviously a gross estimation that does not take into account all the underlying biases that might be present, but it is a starting point for evaluating the sample size.

In addition to what was previously done in other studies [203,210,218,219,227], we decided to widen the set of conditions in order to test our automatic foot strike detection across various circumstances. Therefore, shod running data were recorded at three endurance running speeds. Further, barefoot running data (at preferred speed) were acquired. Moreover, using a treadmill allowed us to have a big number of gait cycles to analyze. This aspect is crucial for a task like running during which a certain amount of adaptation, albeit small, is needed before reaching the cyclic-repetitive state. For this reason, 30 s after at least 60 s accommodation[235] for each condition were recorded, excluding since the beginning the possibility of acquiring only single steps, thus minimizing the effects of artefacts in the assessment of the FSP for each participant and condition.

Since the video analysis (reference technique) is associated with a lack of reliability when dealing with MS cases, only the RS and the joint MFS patterns were used to validate our numerical approach against the reference. The agreement investigation between the two methods produced significantly high values. This will allow us, in the future, to conduct any treadmill-based study by using only our numerical approach for all FSPs (RS, MS, FS and TS). The numerical analysis, supported by a sample size of 145 participants, gave results that are consistent with previous findings[203,226]. There is a clear dominance of RS patterns in the shod condition and MFS patterns in the barefoot condition. These numbers should be interpreted as referred to a sample of habitually shod runners that had no experience of

barefoot running at the moment of the study. The fully automated process avoids any observer influence, thus producing objective results. The high-throughput nature of the numerical analysis helps to dramatically reduce the computation time and increase the efficiency of the FSP assessment.

With our approach, we introduced a foot length correction, which is key for producing accurate results. Peculiar cases can cause difficulties in analyzing data. The TS, for instance, produces a much shorter footprint than the original shoe (or foot). This would lead, without any additional information about the real length of the shoe or foot, to a wrong output. The algorithm would consider the footprint as complete even if only the forefoot and a portion of the midfoot (typical TS case) formed it. To avoid these singularities, the automatic algorithm needs the shoe (or foot) length as an initial input. A quick and easy measurement of the shoe length (for the shod condition analysis) and of the foot (for the barefoot cases) allows our algorithm to correctly locate the pressure information inside the real footprint. Therefore, every possible special case can be automatically analyzed. Furthermore, this method allows for within-person analysis, taking into account any possible asymmetries. This feature would permit a higher level of online-feedback, increasing the amount of available information for both the researcher and the participant.

A potential limitation of this validation might be in its specificity to treadmill running. Even if there is evidence of similarity between overground and treadmill running[145], most of the participants in this study run predominantly outdoors rather than on a treadmill. Also, the participants chose their own footwear and speed, thus not allowing for generalized conclusions on these parameters. Another possible limitation is undoubtedly given by the measurement system. The big size of the sensors (8.47 x 8.47 mm) and the relatively low sampling rate (120 Hz) of the pressure plate, may lead to accuracy issues when the requirements on the measurement uncertainty are particularly stringent.

3 Second study – On the Methodological Implications of Extracting Muscle Synergies from Human Locomotion

Alessandro Santuz[1,2], Antonis Ekizos[1,2], Lars Janshen[1], Vasilios Baltzopoulos[2,3], Adamantios Arampatzis[1,2]

[1]Department of Training and Movement Sciences, Humboldt-Universität zu Berlin, Berlin, Germany
[2]Berlin School of Movement Science, Humboldt-Universität zu Berlin, Berlin, Germany
[3]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, Great Britain

Electronic version of the article published as International Journal of Neural Systems, 2017, 27 (5), 1750007, https://dx.doi.org/10.1142/S0129065717500071 © World Scientific Publishing Co Pte Ltd, https://www.worldscientific.com/worldscinet/ijns.

3.1 Abstract

We investigated the influence of three different high-pass (HP) and low-pass (LP) filtering conditions and a Gaussian (GNMF) and inverse-Gaussian (IGNMF) non-negative matrix factorization algorithm on the extraction of muscle synergies from myoelectric signals during human walking and running. To evaluate the effects of signal recording and processing on the outcomes, we analysed the intraday and interday computation reliability. Results show that the IGNMF achieved a significantly higher reconstruction quality and on average needs one less synergy to sufficiently reconstruct the original signals compared to the GNMF. For both factorizations, the HP with a cut-off frequency of 250 Hz significantly reduces the number of synergies. We identified the filter configuration of 4th order, HP 50 Hz and LP 20 Hz as the most suitable to minimize the combination of fundamental synergies, providing a higher reliability across all filtering conditions even if HP 250 Hz is excluded. Defining a fundamental synergy as a single-peaked activation pattern, for walking and running we identified five and six fundamental synergies respectively using both algorithms. The variability in combined synergies produced by different filtering conditions and factorization methods on the same data set suggests caution when attributing a neurophysiological nature to the combined synergies.

3.2 Introduction

Since the theory of muscle synergies was proposed[9], it has been generally accepted that the central nervous system (CNS) can simplify the production of movements.[52,241] This goal might be achieved by reducing the degrees of freedom through a linear combination[241] of specific muscle activation patterns, called synergies.[52] It has been demonstrated that the two most common types of human locomotion (walking and running) are likely to be controlled by shared synergies.[91] Several studies showed that synergies reside in the brainstem or spinal cord and follow a modular organization.[52,58,68,242–245] These low dimensional units, via descending or afferent pathways, produce a complex electromyographic (EMG) pattern in muscles,[58,242] creating a locomotor drive mediated by a certain amount of supraspinal control.[244,246] Synergies similar to those found at a spinal or muscular level can be observed also in the motor cortex of the primate and cat.[135,186,247] This suggests a high level of cooperation within the CNS structure at all levels.

Several unsupervised learning methods, all aiming to reduce the high dimensional EMG input into a small number of synergies using matrix factorization, are available.[74] They all indicate that the synergies observed in the EMG signals could be neurophysiological entities used by the CNS for generating natural motor behaviors.[74,243,247] Principal component analysis, factor analysis, independent component analysis and non-negative matrix factorization (NMF, used in this study) are some of these methods, which have been shown to produce similar results.[74]

Apart from answering fundamental research questions, the idea of using the muscle synergies for injury prevention, diagnosis or rehabilitation is certainly appealing[59], also considering the growing of the brain-computer interface field.[248–251] Several studies already attempted to use the NMF outcomes as neurophysiological markers for gait stability,[113] in post-stroke patients[103,119,252–254] and in patients with spinal cord injury.[255] In the recent past, some studies investigated methodological issues including the influence of the number and choice of muscles, as well as the number of step cycles, on factorization.[92,256] Another study focused entirely on the influence of the NMF factorization algorithm on the results,[78] finding new ways to improve the NMF performance. However, not a lot of attention has been given,

to date, to the reliability of the used methodologies due to the repeatability and reproducibility of the EMG recordings, the effects of EMG pre-processing[257] and the type of NMF algorithm used to extract the synergies. These are principal issues to be investigated in order to better interpret the results from a neurophysiological point of view.

With this study, we aimed to enrich our knowledge about the NMF initial conditions, by comparing EMG pre-processing parameters, types of algorithms and the similarity metrics for the evaluation of results and their reliability. We collected muscle activities during four locomotion conditions: level and incline walking, shod and barefoot running. For reliability analysis purposes, we measured our participants twice, in two different days. Then, we pre-processed the data with several filtering conditions, ran the NMF using two algorithms[76,78] and evaluated the results' reliability using two similarity metrics. The two algorithms differ in their ability to model the EMG signal-dependent noise properties and are based on the Gaussian and inverse-Gaussian distributions.[78] Devarajan and Cheung[78] suggested that signal-dependent noise formulations should model the noise properties of the EMG signal better than Gaussian formulations.[78,258] To understand the tuning of the computational tools used daily is key to broaden the spectrum of future findings.

This work aims to fill the missing information on the afore-mentioned matters, juxtaposing with the few other studies concerning the methodological implications of using the muscle synergies concept.

3.3 Methods

3.3.1 Experimental Design

Twenty healthy, young, adults were recruited (10 male, 10 female, see Table 7). All the participants were regularly active, right dominant and did not use orthotic insoles. All of them gave informed consent for the experimental procedure, according to the rules of the local scientific board. The experimental design was approved by the Ethics Committee of the Humboldt-Universität zu Berlin. None showed or reported any history of neuromuscular or musculoskeletal impairments. Moreover, in the six months prior to the measurements day, none of them had suffered any injury to the lower limbs. The effect of different initial

conditions on the analysis was assessed by looking at several parameters. First, the filtering conditions imposed to the raw EMG data. Second, the algorithms used to factorize the activation signals. Third, the metrics adopted to evaluate similarities.

Table 7 Participants' anthropometric characteristics and speeds. Values are reported as mean ± s.d.

	Total	M	F
n	20	10 (50%)	10 (50%)
Height [cm]	174 ± 9	180 ± 5	169 ± 8
Body mass [kg]	69 ± 12	77 ± 8	60 ± 8
BMI [kg/m^2]	22 ± 2	24 ± 2	21 ± 2
Age [years]	29 ± 6	31 ± 7	28 ± 5
Pref. running speed [m/s]	2.8 ± 0.4	2.9 ± 0.4	2.6 ± 0.2
Pref. walking speed [m/s]	1.4 ± 0.2	1.4 ± 0.2	1.3 ± 0.2

3.3.2 Materials

The EMG activity of 24 ipsilateral muscles (right side) was recorded using one 16-channel (myon m320, myon AG, Schwarzenberg, Switzerland) and one 8-channel (myon RFTD E08, myon AG, Schwarzenberg, Switzerland) wireless surface-EMG systems. The combination of two systems was necessary to provide an adequate number of channels. However, high-density surface EMG devices supporting up to more than 100 channels have been recently developed.[259] The pairs of wet gel Ag/AgCl electrodes for bipolar derivation (N-00-S, Ambu A/S, Denmark, sensor area 15 mm^2) were applied along the main fiber direction of each muscle, at an inter-electrode distance of 22 mm (longitudinal axis). The acquisition frequency was set to 1000 Hz. Vertical ground reaction forces and plantar pressure distributions were recorded at 120 Hz through a pressure plate (FDM-THM-S, zebris Medical GmbH, Isny im Allgäu, Germany) integrated with a treadmill (mercury, H-p-cosmos Sports & Medical GmbH, Nussdorf, Germany). The pressure plate data were acquired using the proprietary software (WinFDM-T v2.5.1, zebris Medical GmbH, Isny im Allgäu, Germany) and then extracted in a raw format for autonomous post-processing using a validated custom algorithm[196] (R v3.2.2, R Foundation for Statistical Computing, R Core Team, Vienna, Austria). Both the EMG device and the plate were synchronized using an analog signal.

3.3.3 Protocol

The participants conducted a self-selected warm-up[216,232] on the treadmill, in order to choose their comfortable running and walking speeds.

For both running and walking the procedure to find the comfortable speed was implemented using the method of limits[233] and conducted wearing sports shoes. The speed was randomly increased with steps of 0.02 to 0.05 m/s at varying time intervals (around 5 to 10 s) until the participant was comfortable with a specific pace. Then the operation was repeated starting from a faster speed and randomly decreasing it as has been done previously. If the comfortable value was not differing by more than 10% from the previous, the average of the two values was taken as the preferred. Otherwise, the whole procedure was iterated.

After being instructed about the protocol, the participants completed four different tasks, in random order: shod running at the preferred running speed (2.8±0.4 m/s), barefoot running at the preferred running speed (2.8±0.4 m/s), shod level walking at the preferred walking speed (1.4±0.2 m/s) and shod incline walking with 10.0° inclination[260] at 85% of the preferred walking speed (1.2±0.2 m/s).

After an accommodation period of maximum 60 s,[235] the activity of the 24 muscles[256] listed in Table 8 was recorded over around 50 gait cycles (49±4)[92] for each condition. The adaptation to the incline walking condition was limited to 30 s to account for the higher mechanical demands and to avoid unnecessary fatigue.

There were no stops between the trials, except the one or two brakes necessary to change the shoe condition before and after running barefoot. The same randomized protocol was repeated after 15 minutes of rest, without removing the electrodes. After at least 48 h (137±92 h), the routine was replicated placing new electrodes. During this time, the participants were asked not to change their daily routine and not to have any hard training sessions the day prior to the measurements.

Table 8 Muscles analyzed through the EMG devices (ipsilateral, right side of the body). Unless specified differently, the electrodes were positioned on the middle of muscle belly, along the main direction of the fibers. The specifications follow the SENIAM (Surface EMG for non-invasive assessment of muscles) recommendations.

Upper limb	Arm	Shoulder	Pectoral Girdle	Latissimus Dorsi[a]
				Deltoid (ant.)[b]
				Deltoid (post.)[c]
				Trapezius desc. (sup.)[d]
			Upper arm	Biceps Brachii
				Triceps Brachii (long)
Back				Erector Spinæ (long.)[e]
				Splenius Capitis
Abdomen				Rectus Abdominis[f]
				Abdominal Ext. Obl.[g]
	Leg		Upper leg	Semitendinosus
				Biceps Femoris (long)
				Vastus Lateralis
				Vastus Medialis
				Rectus Femoris
				Tensor Fasciæ Latæ[h]
				Adductor Longus
				Gluteus Medius[i]
				Gluteus Maximus[j]
Lower limb			Lower leg	Tibialis Anterior
				Gastrocnemius Med.
				Gastrocnemius Lat.
				Soleus[k]
				Peroneus Longus

[a] One finger width under inferior angle of scapula.

[b] One finger width distal and anterior to acromion.

[c] Two fingers width behind angle of acromion.

[d] Middle of line between acromion and vertebra C7.

[e] Two finger width lateral to vertebra L1.

[f] Two finger width lateral to umbilicus.

[g] 15 cm lateral to umbilicus.

[h] Line from anterior spina iliaca superior to lateral femoral condyle in the proximal 1/6.

[i] Middle of line between iliac crest and greater trochanter.

[j] Middle of line between sacral vertebræ and greater trochanter.

[k] At 2/3 of line between medial condyle of femur to medial malleolus.

3.3.4 Analysis

Surface EMG. The gait cycles' breakdown (cycle identification and segmentation of stance and swing phases) was obtained from the pressure plate raw data. The wireless EMG systems had a built-in band-pass filter (5-500 Hz, 3dB/oct, 4th order). We then applied to the raw EMG signals a high-pass filter (1st, 2nd and 4th order, 20, 50 and 250 Hz cut-off frequencies), a full-wave rectification and then a low-pass filter for the creation of a linear envelope (1st, 2nd and 4th order, 5, 10 and 20 Hz), thus creating 27 filtered data sets, since the order of the high- and low- pass filters was kept constant within a single filtering setup. In addition, a non-filtered data set was produced by not applying any of the previous filters (apart from the built-in band-pass). All filtering methods were of the Infinite Impulse Response Butterworth zero-phase type. Each gait cycle of the 28 data sets was then normalized to 200 data points.[261] It is important to notice, though, that the downsampling and subsequent averaging of the non-filtered data set unavoidably creates a linear envelope of the signal. However, this setting is not influencing the subsequent factorization process, since the objective is not to investigate the physiological correlations between EMG signal and frequency domain. The circa 50 cycles available for each trial were combined into one through averaging.[92] The amplitude was normalized to the maximum activation of the day across all trials and conditions.[52,78,110,262] The EMG data were prepared using R v3.2.2 (R Found. for Stat. Comp.).

Motor modules and motor primitives extraction. Muscle synergies were extracted through a custom algorithm (R version 3.2.2, R Found. for Stat. Comp.) using the NMF approach. We implemented two different NMF update rules for extracting motor modules. The first (GNMF) is based on maximizing the Gaussian likelihood of reconstructing the original EMG signal.[56,76] The second (IGNMF) uses a special case of the generalized Inverse-Gaussian distribution[78] and can be applied when the data exhibit signal-dependent noise, a justified assumption when dealing with EMG signals.[73,79] The muscle activities $V(t)$ recorded from m muscles and normalized to n data points, were represented in an $m \times n$ matrix, factorized as $V(t) \approx V_R(t) = WH(t)$. H is the motor primitive matrix[90] which contains the time-dependent coefficients of the factorization; its dimensions are $r \times n$, where r represents the number of synergies necessary to reconstruct the signal into a matrix V_R. W is the motor modules matrix[253] with dimensions $m \times r$ and it contains the time-invariant muscle weightings. The combination of H and W describes the synergies necessary to accomplish a movement. After

randomly generating two non-negative matrices H and W (values between 0 and 1, uniformly distributed), both algorithms try to converge by maximizing the proportion of explained variation R^2 between the original matrix V and the reconstructed matrix V_R. The GNMF update rules for H and W are presented in Equation 7 and Equation 8. The IGNMF rules in Equation 9 and Equation 10.

$$\begin{cases} H = H\dfrac{(W^T V)}{(W^T W H)} & (7) \\ W = W\dfrac{(V H^T)}{(W H H^T)} & (8) \end{cases}$$

$$\begin{cases} H = H\sqrt{\dfrac{W^T \dfrac{V}{(WH)^2}}{\dfrac{W^T}{V}}} & (9) \\ W = W\sqrt{\dfrac{V \dfrac{H^T}{(WH)^2}}{\dfrac{H^T}{V}}} & (10) \end{cases}$$

The R^2 is expressed by (1-*RSS/SST*), where RSS is the residual sum of squares and SST is the total sum of squares. As already mentioned in the literature,[78] the GNMF and IGNMF algorithms have slightly different formulations of the R^2, since the likelihood distribution of reference is not the same, and can be interpreted in terms of information content of the data. Specifically, the two R^2 quantities have been calculated as shown in Equation 11 and Equation 12.

$$\begin{cases} {R^2}_{GNMF} = 1 - \dfrac{(V - V_R)^2}{(V - \bar{V})^2} & (11) \\ {R^2}_{IGNMF} = 1 - \dfrac{\dfrac{(V - V_R)^2}{VV_R}}{\dfrac{(V - \bar{V})^2}{V^2}} & (12) \end{cases}$$

The limit of convergence was reached when a change in the calculated R^2 between V and V_R was smaller than 0.01% in the last 20 iterations.[79] This was done for a number of synergies successively increased from 1 to 10. The computation was repeated 10 times for each synergy, each time creating new randomized initial matrices, in order to avoid local minima.[80] The solution with the highest R^2 was then selected.

To choose the minimum number of synergies required to represent the original signals, we applied two distinct methods. One is based on the cross-validation of the R^2 values used for describing the reconstruction quality.[79] The curve R^2 vs. synergies is fitted using a simple linear regression model, using all 10 synergies. The mean squared error is then calculated. The same calculation is repeated, each time removing the lower synergy point, until only two points are left or until the error falls below 10^{-5}. To increase the objectivity of this arbitrary threshold, the second method for choosing the minimum number of synergies uses the global minimum of the Akaike Information Criterion (AIC)[78,263] vs. the synergies curve. For a specific algorithm and factorization rank r, the AIC is defined as

$$AIC = 2(RSS + \psi) \qquad (13)$$

where RSS is the residual sum of squares component of the R^2 and $\psi = (m+n)r$ is the total number of parameters for the $m \times n$ input matrix V. The advantage of this method is the objectivity in choosing the best number of synergies without over-fitting.[264]

It might happen that the factorization produces synergies that can be modelled as a linear combination[252] of other, simpler, synergies. We called these simple entities “fundamental”. A fundamental synergy can be defined as an activation pattern whose motor primitive shows a single peak of activation. When two (or more) fundamental synergies are blended into one, a combined synergy appears, like shown in Figure 13. The recognition can be done by manual selection of the fundamental primitives associated to a specific synergy. Due to the great amount of data produced by the different calculated data sets, we implemented a learning algorithm based on a curve-fitting model. The first implementation step consists in choosing some examples of single-peak activation patterns, which might represent a fundamental primitive. The code is then fed by these manually-picked examples of fundamental primitives and a search of similar shapes is done across the whole dataset of

factorized curves. With a first iteration, the primitives that have a high similarity ($R^2 > 0.95$) with the ones present in the manually-created database are added to the set. The number of fundamental primitives is then selected by looking at the motor modules and merging possible repetitions. After updating the database, the code starts the recognition across the entire dataset searching, synergy-by-synergy, for similar primitives (we found $R^2 > 0.5$ to be a good threshold in this phase). Non-recognized curves can then be visually inspected with an interactive routine or automatically identified as new fundamental or combined primitives. This approach, validated in a pilot study, can reproduce the results of a completely manual selection of the curves with a margin of error of ± 5%.

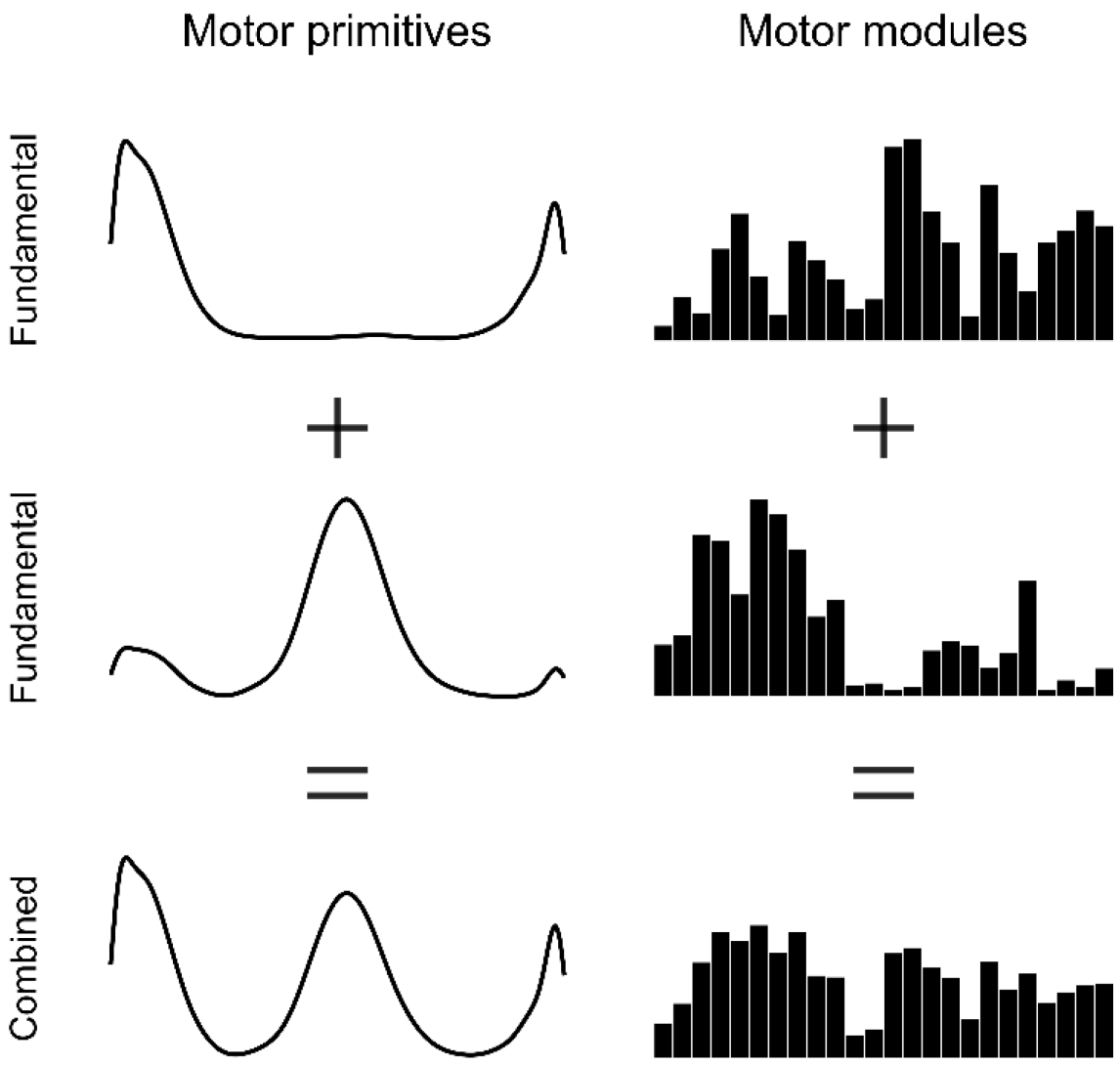

Figure 13 Example of two fundamental synergies combined into one. The production of combined synergies is associated with the factorization method and in turn leads to a reduced number of fundamental synergies.

3.3.5 Statistics

The performance of the two algorithms (GNMF and IGNMF) was assessed by comparing both the proportion of variation in the data explained by each model (i.e. the reconstruction quality measured with R^2) and the computation times. Please note, as mentioned before, that R^2 is dependent on the algorithm used.[78]

We evaluated the different data sets outputs by looking at the similarity between synergies. In this way, the repeatability of the computation and the influences of different filtering methods and factorization algorithms could be assessed. To quantify the similarity of the motor primitives and the motor modules we used two metrics: the coefficient of determination (R^2) and the cosine similarity ($\cos_{sim}$). The first, as previously mentioned, is based on the ratio between the residuals and total sum of squares and is defined as

$$R^2 = 1 - \frac{(H_1 - H_2)^2}{(H_1 - \overline{H_1})^2} \qquad (14)$$

for any two matrices H_1 and H_2 of equal dimensions to be compared. The second is the ratio between the scalar product and the product of the Euclidean norms of the two vectors tested for similarity (i.e. the cosine of the angle between them)[79,80] and is given by the following:

$$\cos_{sim}(a, b) = \cos\theta_{ab} = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|\|\vec{b}\|} \qquad (15)$$

These two metrics have different codomains, but both indicate perfect similarity when equal to 1 (R^2: $\mathbb{R} \rightarrow (-\infty, 1]$, $\cos_{sim}$: $\mathbb{R} \rightarrow [0, 1]$). Using an ANOVA for repeated measures, the similarities of motor primitives and motor modules calculated with the two methods were compared. These metrics can be used for reliability assessment purposes, but they are clearly based on different evaluation scales. The intraday values are obtained by comparing, synergy by synergy, the two within-day trials and then averaging across all the locomotion conditions and all synergies. The interday values are found following the same procedure but comparing the average of the first day with the average of the second day's outcomes. The influence of the filtering conditions on the factorization output has been analyzed at first through a three-way ANOVA for repeated measures applied to all trials of every condition. As dependent

variable, we chose the minimum number of synergies needed by each of the two algorithms (GNMF and IGNMF) to reconstruct the original signal using both the R^2 and the *AIC* criteria. The independent variables (factors) were the filter order and the cut-off frequencies of the high-pass (HP) and low-pass (LP) filters. Each factor consisted of three levels, as mentioned above. All the significance levels were set to $\alpha = 0.05$. If the normality assumptions necessary for the validity of the ANOVA were not satisfied, alternative non-parametric tests were used. The statistical analyses were conducted using R v3.2.2 (R Found. for Stat. Comp.).

3.4 Results

3.4.1 NMF algorithms

The actual number of fundamental synergies recognizable in the factorization results is the same using any of the two NMF algorithms (i.e. 5 for walking and 6 for running). The typical NMF outputs for a specific filtering condition are reported in Figure 14 (level walking), Figure 15 (incline walking), Figure 16 (shod running) and Figure 17 (barefoot running). Here, as a representative case, a 4th order IIR Butterworth zero-phase filter with 50 Hz HP and 20 Hz LP (for linear envelope) cut-off frequencies has been used.

The blending of two or more fundamental synergies into one combined synergy actually reduces the number of recognized fundamental synergies, as shown by the percentages in the four output figures. These numbers represent the amount of recognized fundamental primitives among each condition's trials and are dependent on the factorization method used ($p<0.001$).

Comparing the two NMF algorithms, the minimum number of synergies necessary to reconstruct the original signals is on average five for the GNMF and four for the IGNMF method (Table 9). These results represent the mean values across the filtered and non-filtered data sets and were obtained by employing the R^2 criterion. Using the *AIC* method, though, produced the same results only for the IGNMF algorithm, while for the GNMF created in all cases a minimum at 1 (the same behavior has already been reported by Devarajan and Cheung[78]). The minimum number of synergies was significantly lower in the level compared to the incline walking condition, independently of the factorization method. The largest

difference in the frequency of appearance of fundamental synergies was observed for the fourth synergy. On this matter, no differences could be noted between the two running conditions.

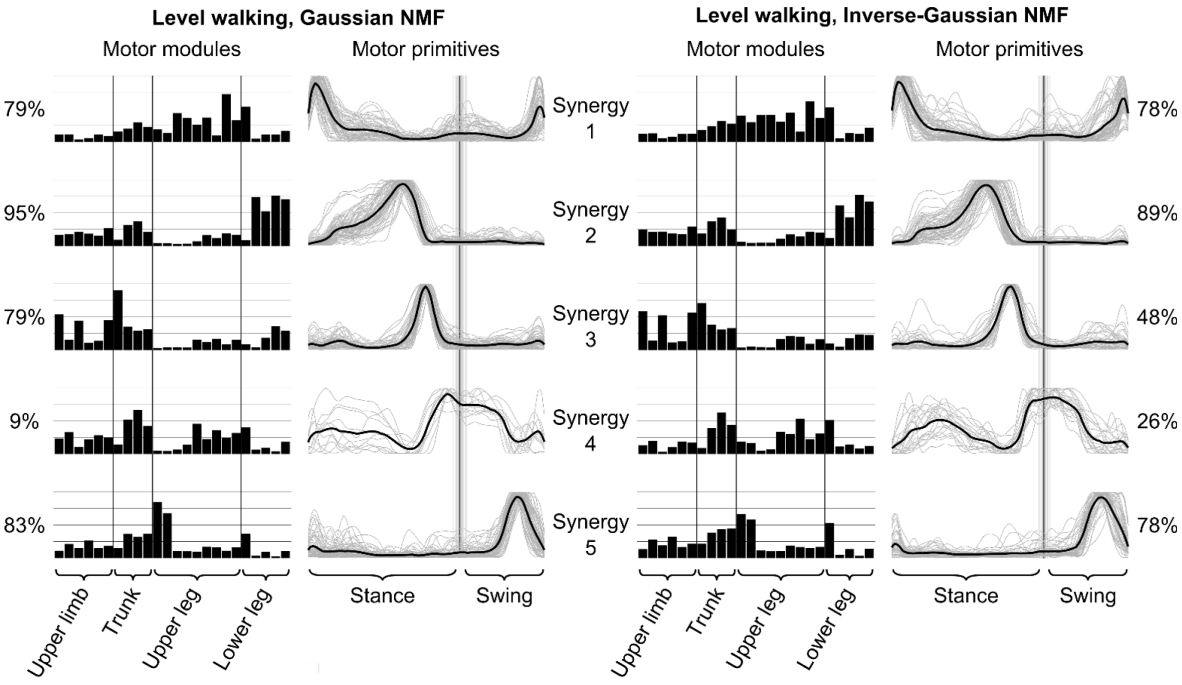

Figure 14 Motor modules and motor primitives obtained with the two factorization methods for level walking. The modules are related, from left to right, to the muscles presented in Table 8, from the top to the bottom and are here represented as mean values of all trials. The mean primitives are represented with a thick black line, while all the recognized trials are denoted by thin grey lines. The x-axis full scale represents one gait cycle, the y-axis the normalized amplitude. Next to each synergy, the percentage of recognized fundamental primitives among all trials is reported. These results were obtained filtering the original signal using a 4th order IIR Butterworth zero-phase filter, with cut-off frequencies high-pass 50 Hz, fully rectifying and then applying a LP 20 Hz IIR Butterworth zero-phase filter for the linear envelope.

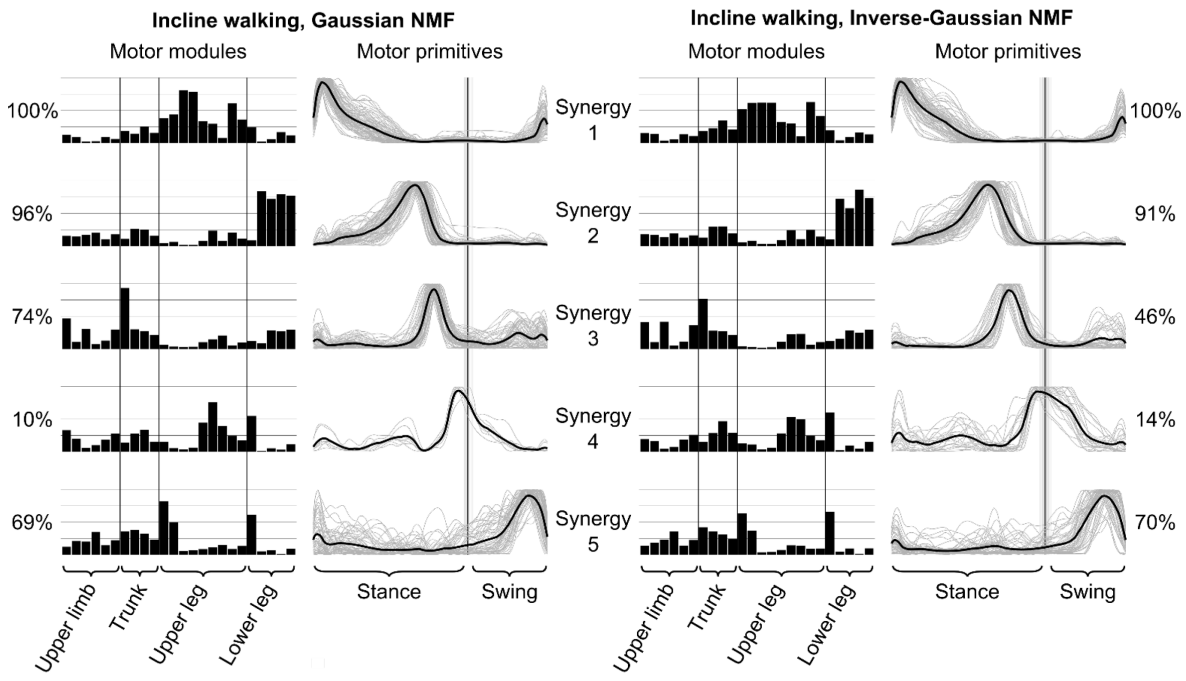

Figure 15 Motor modules and motor primitives obtained with the two factorization methods for incline walking. See Figure 14 legend for further information.

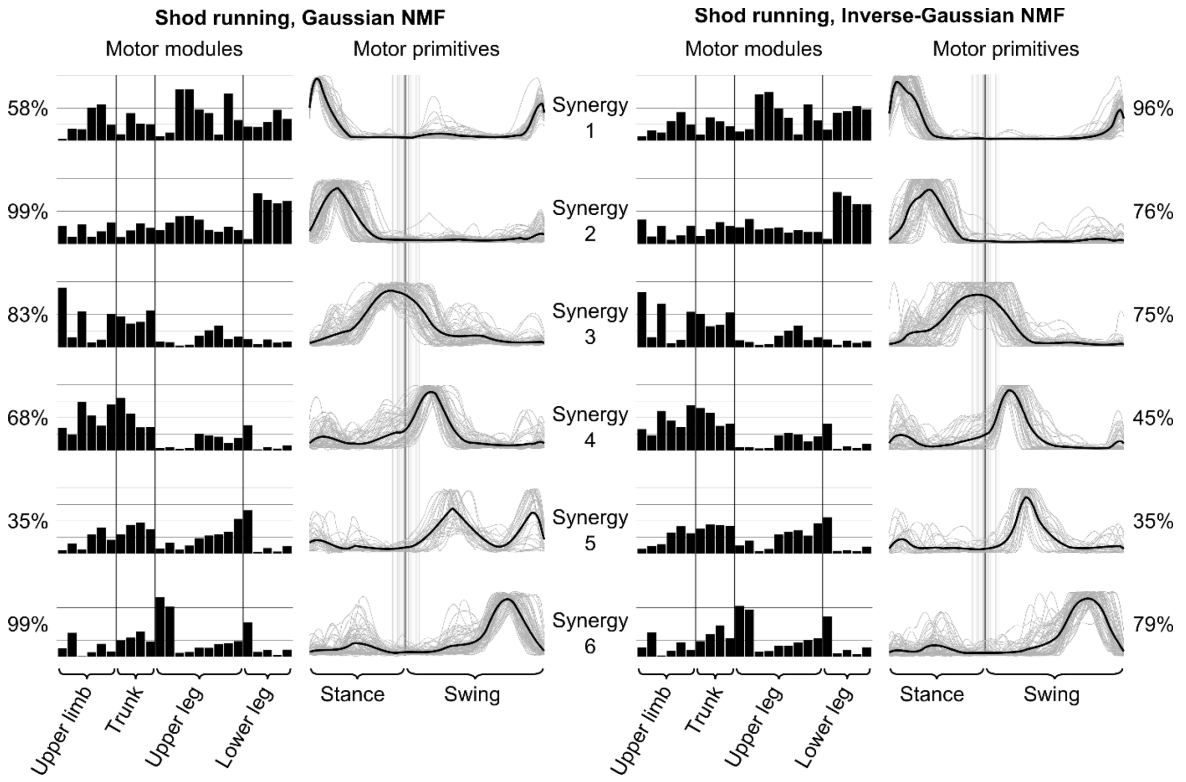

Figure 16 Motor modules and motor primitives obtained with the two factorization methods for shod running. See Figure 14 legend for further information.

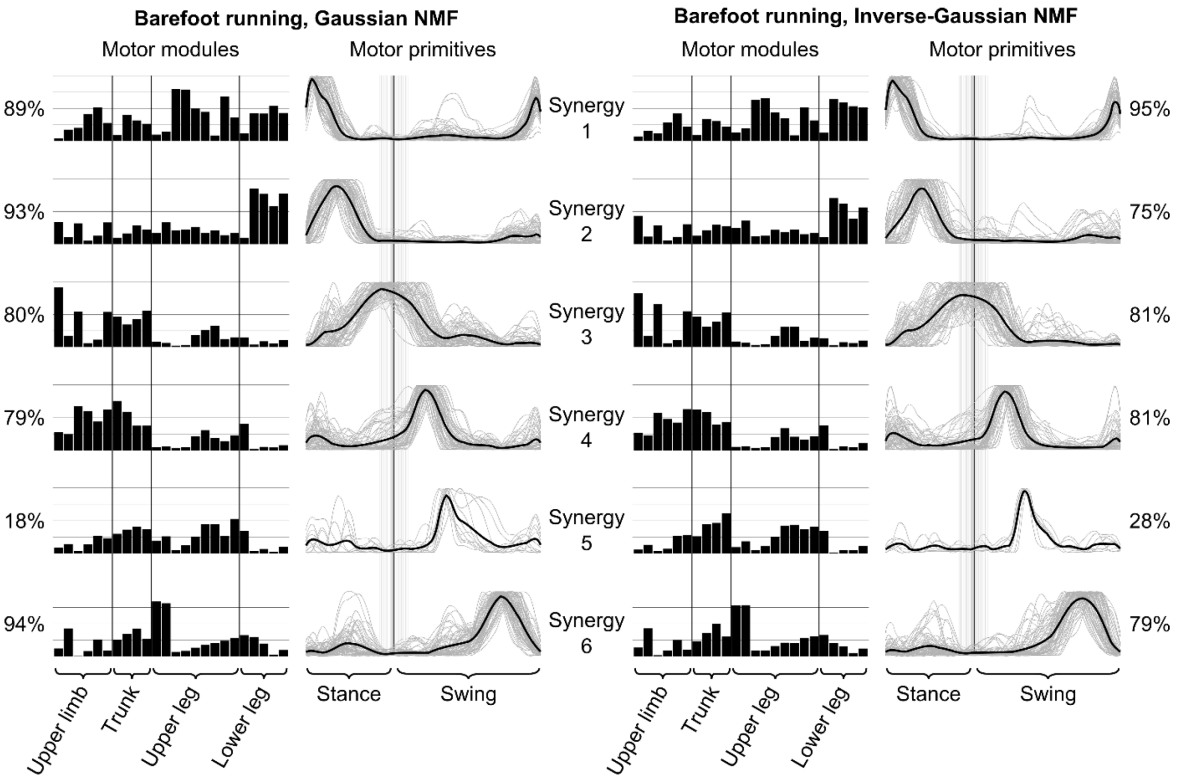

Figure 17 Motor modules and motor primitives obtained with the two factorization methods for barefoot running. See Figure 14 legend for further information.

Table 9 Minimum number of synergies necessary to reconstruct the original set of signals across all the filtering conditions (*p<0.001). GNMF=Gaussian factorization, IGNMF=Inverse-Gaussian factorization.

Condition	GNMF	IGNMF
Walking, level*	4 ± 1.0	3 ± 1.1
Walking, incline*	5 ± 0.8	4 ± 0.9
Running, shod*	5 ± 0.6	4 ± 0.9
Running, barefoot*	5 ± 0.6	4 ± 0.9

The proportion of explained variation R^2 (obtained as described in the methods section using an increasing number of synergies in the model) was independent of the investigated condition. The average values, proven to be significantly different ($p<0.001$), were 0.94 ± 0.02 for the GNMF and 0.99 ± 0.01 for the IGNMF algorithm. This metric can be used to

assess the ability of each algorithm to reconstruct the original signal. Another metric useful for determining the differences in performance is the computation time. Using an Intel® Xeon® X5650 @ 2.66 GHz with 48 GB RAM on Windows 7 64-bit, the computation times for walking and running were significantly lower for the GNMF (1.4 s for every recorded second in both walking and running) compared to the IGNMF (2.0 s for walking and 1.7 s for running). Differences are significant ($p<0.001$).

3.4.2 Filtering conditions

Figure 18 shows the influence of the filtering conditions on the minimum number of synergies needed by the GNMF and IGNMF models to reconstruct the original signal. For both methods, the only consistently significant factor affecting the number of synergies is the HP frequency ($p<0.001$; Figure 18). Other influencing factors are the LP frequency in the GNMF ($p=0.026$) and the filter order in the IGNMF results ($p=0.001$).

A Tukey HSD (honest significant differences) *post-hoc* test reveals the specific role played by the 250 Hz cut-off frequency in the HP setting. This initial condition is the only one that affects both the GNMF ($p=0.04$) and IGNMF ($p=0.02$) results.

The choice of the filter order, on the contrary, only affects the IGNMF approach when comparing the 1st and 4th order-related means ($p=0.001$).

However, these statistics do not take into consideration the number of fundamental synergies that are actually recognized when using a predetermined filtering condition. When including this factor in the analysis, the filtering setup that produces a higher number of fundamental synergies without combining them, is the afore-mentioned 4th order, HP 50 Hz, LP 20 Hz for both running and walking ($p<0.001$).

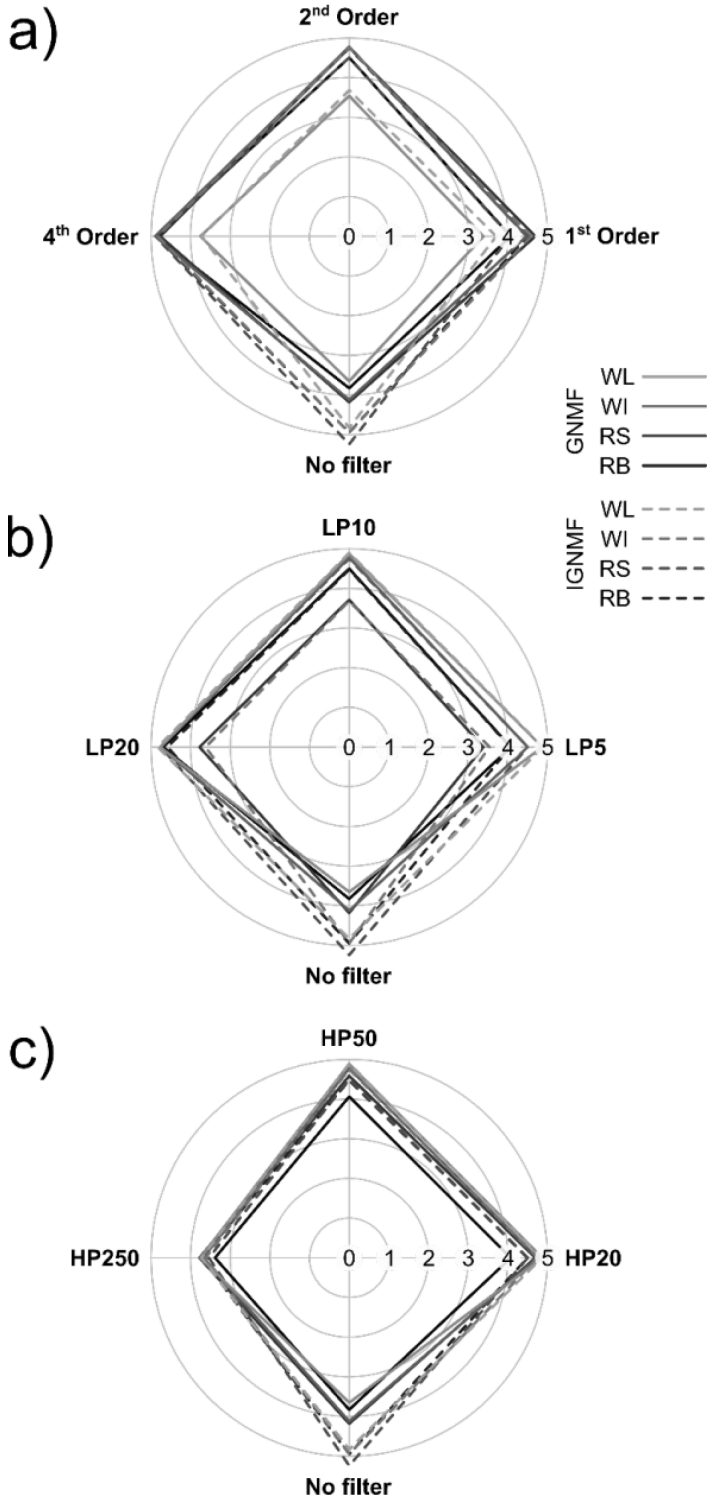

Figure 18 Radar charts showing the influence of the filtering conditions on the minimum number of synergies needed by the two models (GNMF and IGNMF) to reconstruct the original signal. The charts are the result of a Tukey HSD post-hoc test. Plot A: filter order influence. Plot B: low pass (LP) cut-off frequency influence. Plot C: high pass (HP) cut-off frequency influence. GNMF=Gaussian factorization, IGNMF=Inverse-Gaussian factorization, WL=level walking, WI=incline walking, RS=shod running, RB=barefoot running.

3.4.3 Reliability

The coefficient of determination (R^2) and the cosine similarity (cos_{sim}) were the two metrics used to quantify the similarity of the motor primitives and the motor modules belonging to the relative fundamental synergy. The reliability of the results can be assessed by looking at these two numbers. Averaging the outcomes of all filtering conditions, the similarity scenario is the one shown in Table 10. The R^2 values range from 0.76 to 0.86 for motor primitives and from 0.24 to 0.77 for motor modules (intra- and interday, both algorithms). The cos_{sim} values range from 0.94 to 0.97 for motor primitives and from 0.83 to 0.95 for motor modules (intra- and interday, both algorithms).

Table 10 Mean values and relative standard deviations for the similarity metrics R^2 and cos_{sim} across all filtering conditions. The intraday comparison assesses the similarities between two trials of the same day, the interday between different days. GNMF=Gaussian factorization, IGNMF=Inverse-Gaussian factorization, R^2=coefficient of determination, cos_{sim}=cosine similarity.

Comparison	Algorithm	Motor primitives		Motor modules	
		R^2	cos_{sim}	R^2	cos_{sim}
Intraday	GNMF	0.86 ± 0.02	0.97 ± 0.01	0.77 ± 0.02	0.95 ± 0.01
	IGNMF	0.83 ± 0.06	0.96 ± 0.02	0.65 ± 0.09	0.92 ± 0.03
Interday	GNMF	0.79 ± 0.04	0.95 ± 0.01	0.41 ± 0.04	0.86 ± 0.01
	IGNMF	0.76 ± 0.06	0.94 ± 0.02	0.24 ± 0.07	0.83 ± 0.03

It is possible to identify the filtering parameters that give the highest reliability values by removing the data sets filtered with HP 250 Hz and conducting another three-way ANOVA for repeated measures. We considered as reliability variables the similarity metrics R^2 and cos_{sim} after the interday and intraday evaluation. After a Tukey HSD *post-hoc* test, only three filtering conditions gave significant ($p<0.05$) negative mean differences. These three configurations provided significantly higher reliability values than all other filtering setups (in average they show an increase from 1.1 to 1.9 % of the values reported in Table 10). The

three configurations are the non-filtered one and the 1st and 2nd order filter with HP and LP cut-off frequencies of 50 and 5 Hz, respectively. The best configuration to minimize the production of combined synergies (4th order, HP 50 Hz, LP 20 Hz) produces also reliability values higher than the values of Table 10, with an average increase in the similarity values of 1.7 %, but with a higher standard deviation.

3.5 Discussion

3.5.1 NMF algorithms

Five synergies for walking and six for running were sufficient to account for variability of muscle activation using both algorithms. We identified these synergies as fundamental and defined them as activation patterns whose motor primitives show a single peak of activation. However, our idea about fundamental synergies is not purely mathematical. We identify as fundamental a synergy that contains the simplest activation pattern possible. This is also the reason why we refer at this fundamental pattern as motor primitive: there is no possibility of producing such a complex gait movement with a simpler command. From a physiological point of view, we think that this speculation partially agrees with the idea of Bizzi *et al.* about modular organization.[58] According to Bizzi, "the cortex selects and combines the appropriate spinal interneuronal modules, and supplies the modules with temporal patterns of activation appropriate for a behavior execution". We think that, from a system economy point of view, supraspinal-derived activation patterns with only one peak can be more easily stored and, eventually, combined downstream to produce movement.

The two algorithms hereby examined are both based on the concept of decomposing a high-dimensional non-negative matrix into two low-dimensional non-negative matrices, but show different behaviors. Even if the number of fundamental synergies recognizable in the factorized data is the same for both approaches, the order of the resulting factorization matrices is often dissimilar. More specifically, the IGNMF approach has a tendency to reconstruct the original signal with fewer synergies than the GNMF. This is mainly due to the propensity of the IGNMF method to produce more combined synergies than the GNMF. In the present study, we show that different filtering conditions and factorization methods

produce a great variability for combined synergies on the same data set in healthy people. This phenomenon suggests caution when attributing a neurophysiological nature to the combined synergies. It has been speculated that the combination of two or more synergies into one might be related to a reduced independence of neural control signals in post-stroke patients in both lower[119] and upper[252] limbs. Gizzi and colleagues found that post-stroke patients don't show a different number of synergies in the affected arm with respect to healthy controls.[253] Moreover, the same authors highlighted a high similarity between healthy and affected motor primitives, even if the modules appeared to be reshuffled. Other studies[254] did not even find significant differences between the affected and unaffected side, despite observing different EMG signals. Such a varied outcomes scenario might provide evidence that the afore-mentioned point about the different NMF algorithms could be of crucial importance in the interpretation of combined muscle synergies.

Both algorithms needed a higher number of synergies to describe running compared to walking as well as incline compared to level walking. This result might support the idea that movements with higher mechanical and/or stabilization demands may need a more specific set of synergies because of the more distinct and complex neuromuscular requirements.

The reconstruction ability of a factorization algorithm can be assessed by comparing the original data set with the reconstructed one using R^2. The IGNMF showed clearly higher reconstruction capabilities when compared to the classic GNMF ($R^2 = 0.94 \pm 0.02$ for GNMF vs. $R^2 = 0.99 \pm 0.01$ for IGNMF, average values across all conditions). It is difficult, though, to say whether this metric is actually describing the quality of an algorithm's neurophysiological picture of reality. For pure performance-assessment purposes, the computational time for convergence could be of more help. However, no relevant or meaningful (even if significant) differences were observed in the computational times between the GNMF and IGNMF implementations.

Both the GNMF and IGNMF methods can be used to factorize EMG data. However, if the objective of the research is on the fundamental and combined synergies, attention should be paid on the different behavior of the two applications. If the main target is to minimize the production of combinations, the GNMF shall be chosen over the IGNMF. Whether, instead, the focus is not on combinations but, for example, on increasing the reconstruction quality,

the attention should be shifted to the IGNMF. Certainly, further research is needed in order to understand if the combination of synergies is only a computation artefact or an actual behavior of the CNS, used to control movement.

3.5.2 Filtering conditions

Although in most cases the filtering of the EMG signals did not affect significantly the recognition of the fundamental synergies, we found some differences related to the HP cut-off frequency. It has been argued that HP cut-off frequencies of 250 Hz and above would improve computational efficiency and remove electrocardiogram contamination from the signal.[265,266] It has been also suggested that such high cut-off frequencies would highlight the relative contribution of motor units deeper than the surface EMG signal[265] and would better represent the electrophysiological processes occurring at the muscle fiber.[266] When HP filtering the data with such values (i.e. 250 Hz in this study), the factorization produces a high amount of combined synergies. This reduces the possibility of interpreting the results of the two algorithms, both qualitatively and quantitatively. Like already proposed in a recent study,[252] some methods could be used to decompose the combined synergies by looking at them as linear combinations of fundamental ones. This is a promising way of looking at this kind of data. However, the method requires further analysis in order to understand whether the combination is a mere computational product or a result of an actual behavior of the central nervous system. Moreover, an algorithm for the recognition of fundamental synergies that does not rely on any "manual" operation would help in reducing the amount of bias in the whole analysis process.

The order of the filter does not play a big role in the results and might be freely chosen. However, the filtering setup that produces the lowest number of combined synergies is the 4th order, HP 50 Hz, full rectification, linear envelope by LP 20 Hz processing sequence. This is consistent for both algorithms and both locomotion conditions (running and walking). When looking at the afore-mentioned studies analyzing the behavior of synergies in post-stroke patients, the variety of filters choice reflects the heterogeneity of results. Two studies that found differences in the upper[252] and lower[119] limbs pre-processed the EMG data using a HP 50 Hz, LP 20 Hz and a HP 40 Hz, LP 4 Hz setup, respectively (the filter order is hereby omitted since it proved to not influence the results). Coscia and colleagues,[103] who found

differences in primitives but not in modules for the lower limbs, used a LP 10 Hz setup for the linear envelope. On the contrary, the study that reported dissimilarities in modules but not in primitives for upper limbs and trunk,[253] pre-processed the EMG data using also a LP 10 Hz setup for the linear envelope. While the study that found no differences in the upper limbs coordination patterns[254] used a HP 50 Hz, LP 20 Hz filtering setup. Thus, the diversity of some conclusions might be attributed to the signal processing techniques used. When aiming to minimize the amount of combined synergies for both methods (GNMF and IGNMF), the filtering condition we found to be more suitable is a 4th order, HP 50 Hz, LP 20 Hz.

3.5.3 Reliability

The way to assess similarity between motor primitives or motor modules pays indeed the toll of setting an arbitrary threshold. The cosine similarity is often used, but threshold values are much dissimilar across various studies. Several values have been suggested for detecting similarity: $cos_{sim} > 0.60$,[267] $cos_{sim} \geq 0.75$,[252] $cos_{sim} \geq 0.80$[92] or $cos_{sim} \geq 0.90$.[79] With our study, we show that, when analyzing the interday repeatability of the method, the obtained cos_{sim} values are specific for motor primitives and motor modules, other than for algorithms. Moreover, we showed that similarity can be assessed also using the coefficient of determination R^2. This metric clearly has different codomain and outputs than the cos_{sim} ($0.76 < R^2 < 0.86$ for motor primitives and from $0.24 < R^2 < 0.77$ for motor modules vs. $0.94 < cos_{sim} < 0.97$ for motor primitives and $0.83 < cos_{sim} < 0.95$ for motor modules). However, the R^2 appeared to give a more evidence-based output. For instance, when assessing the similarity of a curve and itself mirrored with a vertical symmetry, the cos_{sim} might return a value much greater than zero while the R^2 might be negative (feature that would allow for an easier interpretation of the similarity notion). The same happens with horizontal and central symmetries; all cases that would not be evaluated as similar by simple visual inspection.

Once the issue of setting a threshold is addressed, the reliability analysis shows that intraday values are higher than those for interday repetitions. This can be easily predicted since the interday analysis requires the *ex-novo* placement of electrodes. However, when looking at the cos_{sim} metric, values stay high for both primitives and modules, satisfying the vast majority of similarity thresholds that can be found in literature.[79,92,252,267] The R^2, on the other hand, must be interpreted differently due to the diverse co-domain. Nevertheless, it has

proven useful as much as the cos_{sim} for quantifying the similarity and, thus, reliability between trials for both motor primitives and motor modules. This important information allows declaring the NMF approach described in this study reliable for identifying the fundamental synergies underlying human locomotion. The values found for the intraday reliability could establish a set of similarity thresholds for future reference.

3.6 Conclusion

With this study, we critically analyzed the NMF, one of the most common factorization approaches for extracting muscle synergies. We showed that the choice of NMF algorithm could affect the number of combined synergies calculated. Moreover, we identified those EMG pre-processing conditions (i.e. filtering setup) that might affect the factorized outcomes. We recommend a mild or no pre-processing of the EMG signals. To reduce the number of combined synergies in the outcomes of both methods, without penalizing the reliability, we suggest filtering using a 4^{th} order, HP 50 Hz, fully rectifying and then applying a LP 20 Hz IIR Butterworth zero-phase filter for the linear envelope. Undoubtedly, though, only further research on the matter of combinations of synergies could improve our understanding of their nature. Additionally, we discussed the method's reliability and the metrics to assess it. By measuring the intra- and interday similarities of the recognized fundamental synergies with two metrics (R^2 and cos_{sim}), we found an acceptable reliability for the extraction of the motor primitives and we established the similarity thresholds to be used for future reference. This study was conducted collecting data from healthy participants. Therefore, not all the conclusions regarding the potential issues of using NMF in patients can be entirely confirmed.

3.7 Acknowledgements

The authors are grateful to the participants that always showed great commitment and interest during our measurements.

4 Third study – The Influence of Footwear on the Modular Organization of Running

Alessandro Santuz[1,2], Antonis Ekizos[1,2], Lars Janshen[1], Vasilios Baltzopoulos[2,3], Adamantios Arampatzis[1,2]

[1]Department of Training and Movement Sciences, Humboldt-Universität zu Berlin, Berlin, Germany
[2]Berlin School of Movement Science, Humboldt-Universität zu Berlin, Berlin, Germany
[3]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, Great Britain

This is a post-peer-review, pre-copyedit version of an article published in Frontiers in Physiology, 2017, 8 (958). The final authenticated version is available online at: https://dx.doi.org/10.3389/fphys.2017.00958.

4.1 Abstract

For most of our history, we predominantly ran barefoot or in minimalist shoes. The advent of modern footwear, however, might have introduced alterations in the motor control of running. The present study investigated shod and barefoot running under the perspective of the modular organization of muscle activation, in order to help addressing the neurophysiological factors underlying human locomotion. On a treadmill, 20 young and healthy inexperienced barefoot runners ran shod and barefoot at preferred speed (2.8±0.4 m/s). Fundamental synergies, containing the time-dependent activation coefficients (motor primitives) and the time-invariant muscle weightings (motor modules), were extracted from 24 ipsilateral electromyographic activities using non-negative matrix factorization. In shod running, the average foot strike pattern was a rearfoot strike, while in barefoot running it was a mid-forefoot strike. In both conditions, five fundamental synergies were enough to describe as many gait cycle phases: weight acceptance, propulsion, arm swing, early swing and late swing. We found the motor primitives to be generally shifted earlier in time during the stance-related phases and later in the swing-related ones in barefoot running. The motor primitive describing the propulsion phase was significantly of shorter duration (peculiarity confirmed by the analysis of the

spinal motor output). The arm swing primitive, instead, was significantly wider in the barefoot condition. The motor modules demonstrated analogous organization with some significant differences in the propulsion, arm swing and late swing synergies. Other than to the trivial absence of shoes, the differences might be deputed to the lower ankle gear ratio (and the consequent increased system instability) and to the higher recoil capabilities of the longitudinal foot arch during barefoot compared to shod running.

4.2 Introduction

In the last decade, the study of locomotion in evolutionary anthropology has been increasingly focusing on endurance running. Humans, compared to non-human primates, show exceptional endurance running speeds[202]. However, the advent of modern running shoes is contemporary history compared to the two million-years-old fossil evidence of running as a derived capability of the genus *Homo*[202,204]. Running barefoot or in minimal footwear has been the predominant condition for most of the human history[204]. Hence, it can be expected that the strategies adopted to run barefoot might differ from those employed to run shod.

During running, the foot can strike the ground in multiple ways, called foot strike patterns (FSPs). Rearfoot (RS), midfoot (MS) and forefoot (FS) strike are the common classifications, depending on the location of the first contact area with the ground[203]. We recently found that if almost 90% of the population adopts a RS pattern when running shod, only half maintain it when switching to barefoot[196], changing to either MS or FS (joined in an unique pattern and indicated as mid-forefoot strike, MFS for brevity). Adopting MFS patterns can increase the plantarflexors activity, reduce the ground contact times and affect the kinetics and kinematics of the whole gait cycle[203,204,268–270]. Therefore, we can argue that switching between the two conditions of running shod and barefoot does not only imply kinematic and kinetic changes, but might involve a different organization of movement. From a motor control perspective, this assumption can be investigated by analyzing the modular organization of muscle activity before and after altering the running condition.

Since the late 1960s[9] it has been accepted that the central nervous system can simplify the production of movements by avoiding the activation of each muscle separately[52,241]. This important feature might be implemented by reducing the degrees of freedom through a linear

combination[241] of specific muscle activation patterns, called synergies[52]. A number of studies were able to show that synergies reside in the brainstem or spinal cord and follow a modular organization[52,58,242,244,245]. Recently, a study in mice using optogenetics to isolate the excitatory and inhibitory neuronal populations could show a strong specificity in the spinal cord topography[271]. The synergies as low dimensional units, via descending or afferent pathways, produce a complex electromyographic (EMG) pattern in muscles[58,242], creating a locomotor drive mediated by a certain amount of supraspinal control[244]. During walking, the same amount of basic activation patterns could be found in patients with spinal cord injury and in healthy participants at different speeds and gravitational loads[272]. Synergies similar to those found in humans at a spinal[272] or muscular level can be observed also in the motor cortex of the primate and cat[135,186]. Moreover, studies on the excitability of the corticospinal system showed that training can improve task-specific brain organizations[273–275]. This suggests a high degree of cooperation within the central nervous system's structure at all levels. In this study, we used an unsupervised learning method called non-negative matrix factorization (NMF)[56] for reducing the high dimensional EMG input into a small number of synergies. We focus on the comparison between the modular organization of shod and barefoot running. Compared to the analysis of direct EMG signals, the muscle synergies concept has the clear advantage of being a high-throughput approach for analyzing muscle activities. In fact, it does not only provide the researcher with an automatic, low-dimensional, clustering of the activations during the gait cycle, but it also identifies the weighted contribution of each muscle for producing a certain movement.

The objective of the current study was to investigate the modular organization of shod and barefoot running using muscle synergies in order to gain new knowledge about the neurophysiological factors underlying human locomotion. Based on reported changes in the kinematic, kinetic[204] and EMG[91,268–270,276,277] characteristics of shod and barefoot running, we hypothesized that there is an alteration in the modular organization between the two conditions that might be associated with the specificity of the respective foot strike with the ground. In a first step we calculated the similarities between trials of the same condition using the coefficient of determination (R^2) and set their repeatability (intraday) thresholds[75]. Afterwards, we investigated the similarities between the two conditions and compared with the intraday thresholds. This two-step process allowed an improved objective and quantitative interpretation of the conventionally qualitative output given by methods like the NMF.

4.3 Material and methods

4.3.1 Experimental protocol

Twenty healthy and young adults were recruited (10 male, height 180 ± 5 cm, body mass 77 ± 8 kg, age 31 ± 7 years, 10 female, height 169 ± 8 cm, body mass 60 ± 8 kg, age 28 ± 5 years). All the participants were regularly active and did not use orthotic insoles and did not have any previous experience of barefoot running. None showed or reported any history of neuromuscular or musculoskeletal impairments, or any head or spine injury at the time of the measurements or in the previous six months. This study was reviewed and approved by the Ethics Committee of the Humboldt-Universität zu Berlin. All the participants gave written informed consent for the experimental procedure, in accordance with the Declaration of Helsinki.

The muscle activity of 24 ipsilateral muscles was recorded using one 16-channel (myon m320, myon AG, Schwarzenberg, Switzerland) and one 8-channel (myon RFTD E08, myon AG, Schwarzenberg, Switzerland) wireless surface-EMG systems. The acquisition frequency was set to 1000 Hz (16 ms latency, constant). Vertical ground reaction forces were recorded at 120 Hz through a pressure plate (FDM-THM-S, zebris Medical GmbH, Isny im Allgäu, Germany) integrated with a treadmill (mercury, H-p-cosmos Sports & Medical GmbH, Nussdorf, Germany). The pressure plate data were acquired using the proprietary software (WinFDM-T v2.5.1, zebris Medical GmbH, Isny im Allgäu, Germany) and then extracted in a raw format for autonomous post-processing using a validated custom algorithm[196] written in R version 3.4.1 (R Foundation for Statistical Computing, R Core Team, Vienna, Austria). The EMG devices and the plate were synchronized using an analogue signal.

The participants completed a self-selected warm-up on the treadmill, in order to choose their comfortable shod-running speed. The procedure to find the comfortable speed was implemented using the method of limits[233]. The speed was randomly increased with steps of 0.02 to 0.05 m/s at varying time intervals (around 5 to 10 s) until the participant was comfortable with a specific pace. The operation was then repeated starting from a faster speed and randomly decreasing it as previously done. If the comfortable value did not differ more than 10% from the previous, the average of the two values was taken as the preferred.

Otherwise, the whole procedure was iterated. The warm-up protocol typically lasted between 5 and 10 min. After being instructed about the protocol, the participants completed two different tasks, in random order: shod running at the preferred running speed (2.9 ± 0.4 m/s for male, 2.6 ± 0.2 m/s for female) and barefoot running at the same speed.

For each condition, the muscle activity of the 24 ipsilateral (right side) muscles was recorded: *splenius capitis* (SP), *trapezius* (descending, TR), *latissimus dorsi* (LD), *deltoid* (anterior, DA), *deltoid* (posterior, DP), *biceps brachii* (BB), *triceps brachii* (long head, TB), *erector spinæ* (*longissimus*, L1 vertebra, ES), *rectus abdominis* (RA), *abdominal external oblique* (AE), *gluteus medius* (ME), *gluteus maximus* (MA), *adductor longus* (AL), *tensor fasciæ latæ* (FL), *rectus femoris* (RF), *vastus medialis* (VM), *vastus lateralis* (VL), *semitendinosus* (ST), *biceps femoris* (long head, BF), *tibialis anterior* (TA), *peroneus longus* (PL), *gastrocnemius medialis* (GM), *gastrocnemius lateralis* (GL) and *soleus* (SO). Around 50 gait cycles (49 ± 4)[92] were recorded after an accommodation period of maximum 60 s[235]. Between the trials there was a break necessary to change shoes before and after running barefoot. The same randomized protocol was repeated after 15 minutes of rest for use in the (intraday) repeatability analysis, without removing the electrodes.

4.3.2 Foot strike patterns assessment

For every trial, the FSP and the strike index were calculated using a validated algorithm based on the numerical analysis of foot pressure distribution[196]. As we previously suggested[196], the FSPs have been grouped into two categories rather than three: RS and MFS (including MS and FS patterns). The strike index, as originally defined by Cavanagh and Lafortune[201], was calculated as the distance from the heel to the center of pressure at impact relative to total foot length.

4.3.3 Spinal motor output assessment

For the spinal motor output characterization, we mapped the 24 measured EMG activities onto the estimated rostrocaudal location of alpha-motoneurons (MNs) pools in the segments from the second cervical vertebra (C2) to the second sacral vertebra (S2) of the spinal cord[272,278]. The segments T2, T3 and T4 have been excluded from the analysis since they do not innervate any of the considered muscles. The wireless EMG systems had a built-in band-

pass filter (5-500 Hz, 3 dB/oct, 4th order). The EMG signals were high-pass filtered and then full-wave rectified and low-pass filtered using a 4th order IIR Butterworth zero-phase filter with cut-off frequencies 50 Hz (high-pass) and 20 Hz (low-pass for the linear envelope) using R v3.4.1 (R Found. for Stat. Comp.). The amplitude was normalized to the maximum activation recorded for each participant across all conditions[52,78,262]. Each gait cycle was then time-normalized to 200 points[279], assigning 100 points to the stance and 100 points to the swing phase. The cervical segments (C2 to C8) mainly innervate upper limb and neck muscles. The thoracic segments (T1 to T12) connect to the trunk muscles, while the lumbar (L1 to L5) and sacral (S1 and S2) segments innervate the lower limb muscles. The contribution of each muscle to the total estimated activity of the spinal segments was implemented using the myotomal charts developed by Kendall[280]. This method shows the organization of the efferent MNs network directed to the muscles, assuming a common spinal topography among the investigated participants. Without accounting for size differences in MN pools at each spinal level, the motor output of each spinal segment S_j was estimated using the Equation 16[278]:

$$S_j = \frac{\sum_{i=1}^{m_j} \left(\frac{k_{ji}}{n_i} \times EMG_i \right)}{\sum_{i=1}^{m_j} \left(\frac{k_{ji}}{n_i} \right)} \qquad (16)$$

where m_j are the muscles innervated by each segment, n_i is the number of spinal levels that innervate the i^{th} muscle, k_{ij} is a weighting coefficient specific to each muscle and spinal segment (e.g. $k_{ij} = 1$ or $k_{ij} = 0.5$ if S_j is a major or minor MN source, respectively) and EMG_i is the normalized recorded EMG, specific for each participant and trial[278,280].

4.3.4 Modular organization assessment

The gait cycle breakdown was obtained from the pressure plate's raw data. Using a custom algorithm[196], the touchdown was identified as the first non-zero pressure matrix after the last toe-off. The EMG signals were pre-processed using the filtering and normalization conditions reported above.

Muscle synergies data were extracted through a custom script[75] (R v3.4.1, R Found. for Stat. Comp.) using the classical Gaussian NMF algorithm[56] from the first circa 50 gait cycles of

each acquisition[92]. EMG data were pre-processed using the same filtering conditions reported in the previous paragraph. The $m = 24$ time-dependent muscle activity vectors were grouped in an $m \times n$ matrix V, factorized such that $V \approx V_R = WH$. V_R represents the new reconstructed matrix, which approximates the original matrix. The motor primitives matrix H[75,90] contained the time-dependent coefficients of the factorization with dimensions $r \times n$, where r represents the number of synergies necessary to reconstruct the signal and n the number of data points ($n = 200 \cdot$ number of cycles). The motor modules matrix W[75,253] with dimensions $m \times r$, contained the time-invariant muscle weightings. H and W described the synergies necessary to accomplish a movement. The update rules for H and W are presented in Equation 17 and Equation 18. The limit of convergence was reached when a change in the calculated R^2 between V and V_R was smaller than the 0.01% in the last 20 iterations[75,79]. To choose the minimum number of synergies required to represent the original signals, the curve of R^2 values versus synergies was fitted using a simple linear regression model, using all ten synergies. The mean squared error was then repeatedly calculated[75], each time removing the lower synergy point, until only two points were left or until the mean squared error fell below 10^{-5}.

The gait cycle breakdown was obtained from the pressure plate's raw data. Using a custom algorithm[196], the touchdown was identified as the first non-zero pressure matrix after the last toe-off. The EMG signals were pre-processed using the filtering and normalization conditions reported above.

Muscle synergies data were extracted through a custom script[75] (R v3.4.1, R Found. for Stat. Comp.) using the classical Gaussian NMF algorithm[56] from the first circa 50 gait cycles of each acquisition[92]. EMG data were pre-processed using the same filtering conditions reported in the previous paragraph. The $m = 24$ time-dependent muscle activity vectors were grouped in an $m \times n$ matrix V, factorized such that $V \approx V_R = WH$. V_R represents the new reconstructed matrix, which approximates the original matrix. The motor primitives matrix H [75,90] contained the time-dependent coefficients of the factorization with dimensions $r \times n$, where r represents the number of synergies necessary to reconstruct the signal and n the number of data points ($n = 200 \cdot$ number of cycles). The motor modules matrix W [75,253] with dimensions $m \times r$, contained the time-invariant muscle weightings. H and W described the synergies necessary to

accomplish a movement. The update rules for H and W are presented in Equation 17 and Equation 18. The limit of convergence was reached when a change in the calculated R^2 between V and V_R was smaller than the 0.01% in the last 20 iterations [75,79]. To choose the minimum number of synergies required to represent the original signals, the curve of R^2 values versus synergies was fitted using a simple linear regression model, using all ten synergies. The mean squared error was then repeatedly calculated[75], each time removing the lower synergy point, until only two points were left or until the mean squared error fell below 10^{-5}.

$$\begin{cases} H = H\dfrac{(W^T V)}{(W^T WH)} & (17) \\ W = W\dfrac{(VH^T)}{(WHH^T)} & (18) \end{cases}$$

The aforementioned procedure allowed us to extract fundamental and combined synergies from the raw EMG data. A fundamental synergy can be defined as an activation pattern whose motor primitive shows a single peak of activation [75]. When two or more fundamental synergies are blended into one, a combined synergy appears. Due to the lack of consent in the literature on how to interpret combined synergies, they were excluded from the analysis. An example of combined synergies is reported in Figure 19.

The fundamental synergies recognition was implemented using a custom learning algorithm based on a curve-fitting model[75]. The first implementation step consists in choosing some examples of single-peaked activation patterns, which might represent a fundamental primitive. The code is then provided with this training set and a search of similar shapes is done across the whole dataset of factorized curves. With a first iteration, the primitives that have a high similarity ($R^2 > 0.95$) with the ones present in the training set are added to it. The number of fundamental primitives is then selected by clustering similar motor modules. After updating the training set, the code starts the recognition across the entire dataset searching, synergy-by-synergy, for similar primitives (we found $R^2 > 0.50$ to be a good threshold). Non-recognized curves can then be visually inspected with an interactive routine or automatically identified as new fundamental or combined primitives. This approach, validated in a pilot study, can reproduce the results of a completely manual selection of the curves with a margin of error of ± 5%.

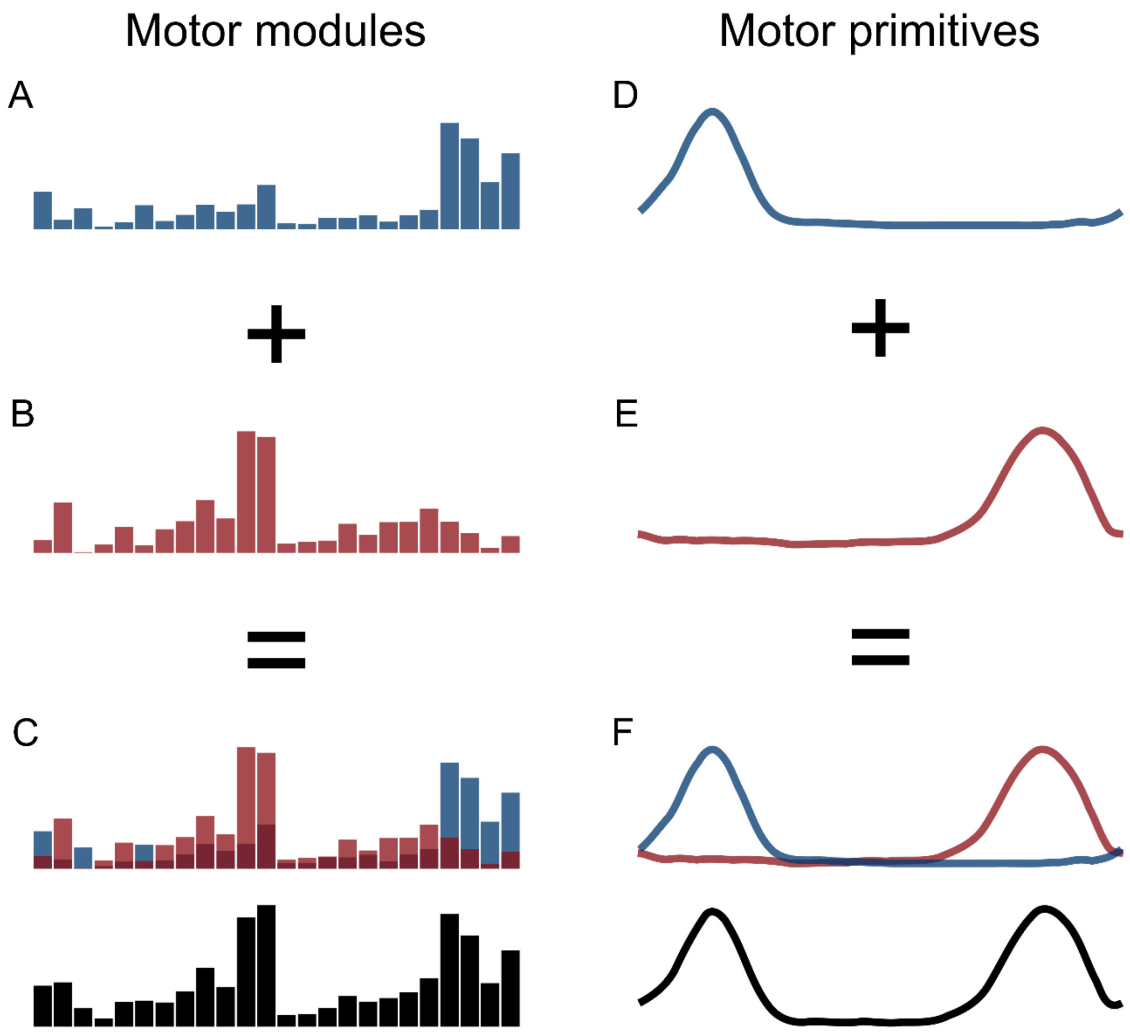

Figure 19 Example of two fundamental synergies combined into one. The histograms in the panels A,B represent the two fundamental sets of motor modules and the curves in the panels D,E the two respective primitives, with arbitrary x- and y-axis units. The combined motor modules and primitives are shown in panels C,F, respectively.

4.3.5 Metrics for comparison of curves

We evaluated the center of activation (*CoA*) and full width half maximum (FWHM) for the resulting curves of the extracted spinal maps and motor primitives (matrix H) in both conditions and types of locomotion. The *CoA* was defined as the angle of the vector (in polar coordinates) that points to the center of mass of that circular distribution[279]. The polar direction represented the gait cycle's phase, with angle $0 \leq \theta_t \leq 2\pi$. The following equations define the *CoA*:

$$\begin{cases} A = \sum_{t=1}^{p} (\cos \theta_t \times P_t) & (19) \\ B = \sum_{t=1}^{p} (\sin \theta_t \times P_t) & (20) \end{cases}$$

$$CoA = \arctan(B/A) \qquad (21)$$

where p is the number of points of each gait cycle (p = 200) and P is the activation vector. The FWHM was calculated as the number of points exceeding each gait cycle's half maximum, after subtracting the gait cycle's minimum[279]. For every trial, both parameters were calculated at each gait cycle and then averaged to proceed with the statistical analysis. A maximum of 50 gait cycles for each acquisition were selected for analysis. The *CoA* and FWHM were analyzed for stance and swing distinctively for spinal maps and over the whole gait cycle for the motor primitives.

4.3.6 Statistics

A two-way ANOVA with repeated measures, followed by a Tukey *post-hoc* analysis with false discovery rate *p-value* adjustment, was used to investigate *CoA* and FWHM between conditions. The same statistics was used for the motor modules, using the muscles and the conditions (shod vs. barefoot) as independent variables. To assess the similarities between the fundamental motor primitives of shod and barefoot running, we used the coefficient of determination R^2. We calculated the similarity values between the pairs of trials recorded during the same day (intraday repeatability) in shod and barefoot running. Then, we compared

them with the similarity values between the two conditions (shod and barefoot running). Type A uncertainty was expressed as $u_A = s/\sqrt{n}$. All the significance levels were set to $\alpha = 0.050$ and the statistical analyses were conducted using R v3.4.1 (R Found. for Stat. Comp.).

4.4 Results

4.4.1 Foot strike patterns and gait parameters

Out of 20 participants, 14 (7 male, 7 female) transitioned from RS (shod) to MFS (barefoot). Three kept the MFS pattern in both conditions and three retained a RS pattern in both shod and barefoot running. The participants demonstrated significant ($p < 0.001$) differences in the average strike index, presenting values of 0.15 ± 0.17 in shod running and 0.53 ± 0.18 in barefoot running (with 0 denoting the most posterior and 1 the most anterior point of the shoe, see Figure 20). Also the average contact times of 301 ± 36 ms and 274 ± 32 ms as well as the average cadence (step frequency) of 162 ± 10 and 166 ± 11 steps/min were significantly different ($p < 0.001$) between shod and barefoot running, respectively. The mean values of the left and right vertical ground reaction forces (VGRFs) normalized to body weight were significantly lower in the barefoot condition (1.82 ± 0.20 for the shod and 1.75 ± 0.16 for the barefoot condition, $p < 0.001$). The impulse (mean values of left and right sides) was significantly lower in the barefoot condition (201 ± 39 N·s versus 186 ± 35 N·s, $p < 0.001$), but the flight time was contrarily higher (70 ± 24 ms for shod and 89 ± 22 ms for barefoot running, $p < 0.001$).

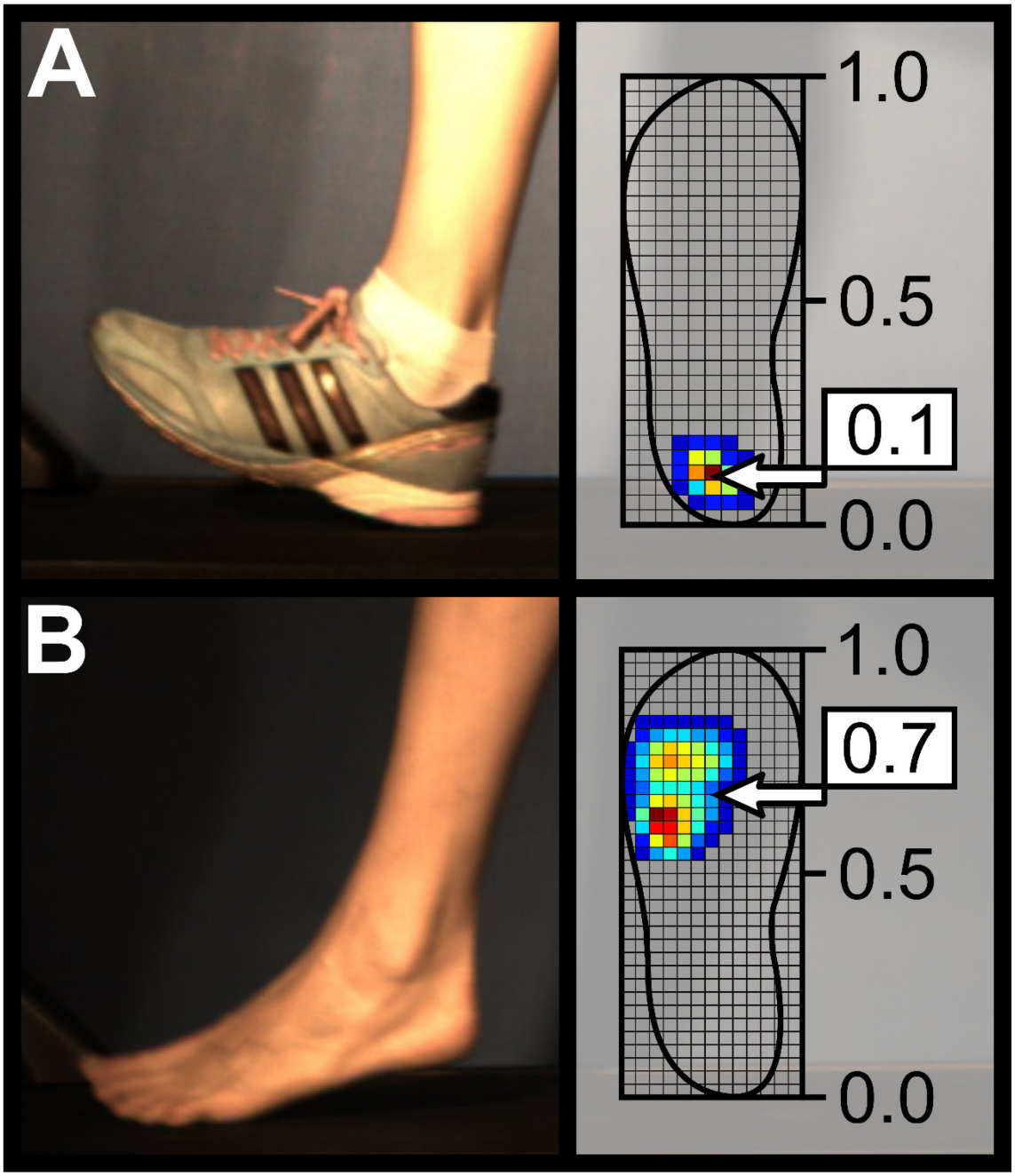

Figure 20 Sagittal view of a typical rearfoot (A) and forefoot (B) strike patterns during shod and barefoot running, respectively. The strike index values extracted from plantar pressure distribution for these two representative cases are presented as well.

4.4.2 Spinal motor output

Figure 21 depicts the average spatiotemporal spinal motor output for shod and barefoot running. The two-way ANOVA identified statistically significant differences in the FWHM of the mapped EMG activities when comparing shod and barefoot running for both the stance (p=0.018) and swing (p=0.019) phase of the gait cycle (Table 11). The *post-hoc* analysis showed significantly lower FWHM in the barefoot condition of segment L4's spinal motor output, innervating the muscles ME, AL, FL, RF, VM, VL, ST, TA and PL. The *CoA* was not significantly different between conditions in neither the stance (p=0.107) or the swing (p=0.091) phase (Table 11).

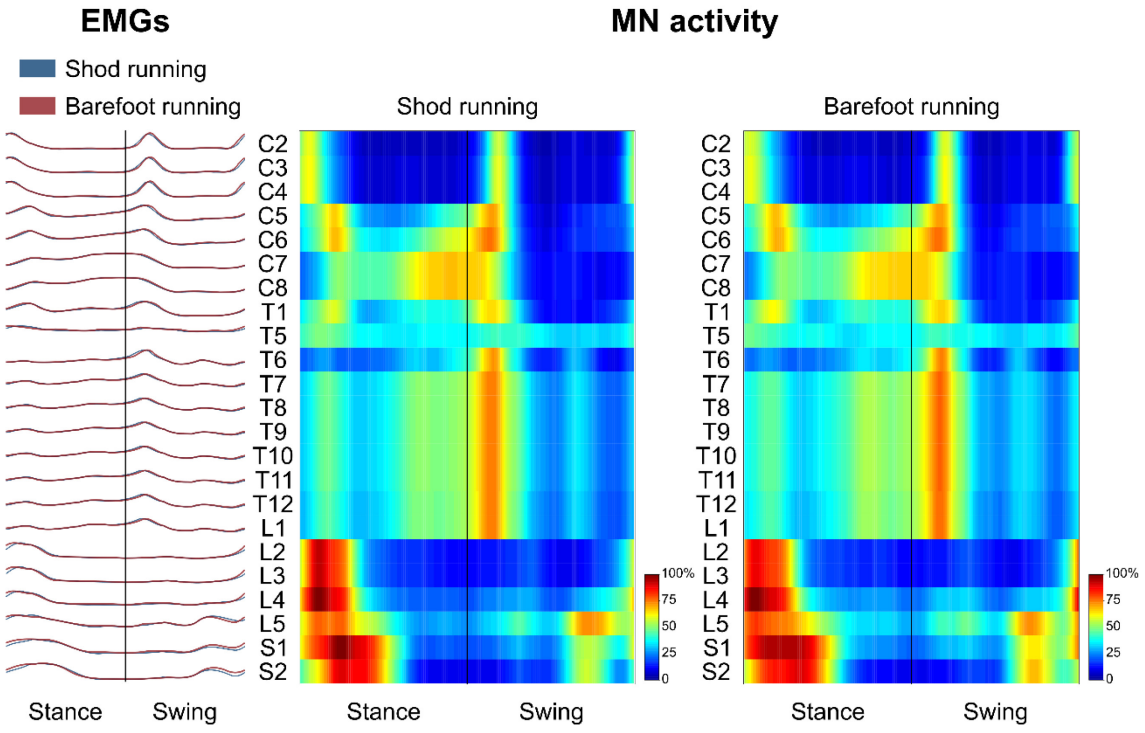

Figure 21 The average spatiotemporal spinal motor output is presented for shod and barefoot running, normalized in amplitude to the maximum of each segment. These curves have been obtained by mapping each of the 24 muscle activations onto the relative spinal segment (cervical from C2 to C8, thoracic from T1 to T12, lumbar from L1 to L5 and sacral from S1 and S2). The two level plots show the average alpha-motoneurons activity for each condition, giving additional information about the absolute activation level (normalization to the maximum of each condition). The stance and swing phases have been temporally normalized to the same amount of data points (100 each). Values are the means across all subjects and all trials.

Table 11 Differences between shod and barefoot running in the center of activity (*CoA*) and full width at half maximum (FWHM) of the electromyographic activities mapped onto the estimated rostrocaudal location of the spinal cord (segments C2 to S2). Results concerning T2, T3 and T4 are not reported since those segments do not innervate any muscle considered in this study. Positive differences ($\Delta_{E,U}>0$) denote bigger values in the barefoot condition, whereas negative differences imply lower values.

	CoA		**FWHM**			
	Stance ***p*=0.107**	**Swing** ***p*=0.091**	**Stance** ***p*=0.018***		**Swing** ***p*=0.019***	
Segment	$\Delta_{S,B}$	$\Delta_{S,B}$	$\Delta_{S,B}$	***p-value***	$\Delta_{S,B}$	***p-value***
C2	+2.6%	+2.7%	-0.8%	0.718	-0.8%	0.712
C3	+2.5%	+2.5%	-1.0%	0.658	-1.0%	0.654
C4	+2.5%	+2.5%	-1.0%	0.658	-1.0%	0.654
C5	-0.2%	-0.2%	-0.9%	0.698	-0.9%	0.669
C6	-0.5%	-0.5%	-0.6%	0.794	-0.6%	0.818
C7	-0.4%	-0.4%	-1.7%	0.362	-1.6%	0.403
C8	-0.4%	-0.4%	-1.7%	0.362	-1.6%	0.403
T1	+0.4%	+0.5%	-1.0%	0.657	-1.1%	0.614
T5	+3.0%	+2.9%	-1.7%	0.371	-1.8%	0.342
T6	-0.9%	-0.9%	+0.2%	0.949	+0.2%	0.959
T7	+0.0%	+0.0%	+0.2%	0.954	+0.2%	0.930
T8	+0.0%	+0.0%	+0.2%	0.954	+0.2%	0.930
T9	+0.0%	+0.0%	+0.2%	0.954	+0.2%	0.930
T10	+0.0%	+0.0%	+0.2%	0.954	+0.2%	0.930
T11	+0.0%	+0.0%	+0.2%	0.954	+0.2%	0.930
T12	-0.2%	-0.2%	0.1%	0.984	0.1%	0.972
L1	-0.2%	-0.2%	0.1%	0.984	0.1%	0.972
L2	+1.7%	+1.9%	-2.3%	0.210	-2.3%	0.212
L3	+1.7%	+1.9%	-2.3%	0.210	-2.3%	0.212
L4	+0.2%	+0.2%	*-3.7%*	*0.036**	*-4.0%*	*0.025**
L5	-0.6%	-0.6%	-2.1%	0.252	-2.3%	0.209
S1	+0.5%	+0.6%	-0.1%	0.995	+0.0%	0.996
S2	+0.6%	+0.6%	1.0%	0.641	1.0%	0.630

4.4.3 Modular organization

The average number of recognized fundamental synergies during running was significantly different between the two conditions (3.9 ± 0.6 for shod and 3.6 ± 0.6 for barefoot running, $p < 0.001$). However, in both conditions, five fundamental activation patterns could be identified (Figure 22 and Figure 23). The five fundamental synergies extracted during both shod and barefoot running, were associated with temporally different phases of the gait cycle and ordered according to the timing of each motor primitive's global maximum (Figure 22 and Figure 23). The first synergy (peak at ~8% of the stance phase) functionally referred to the body weight acceptance, with a major involvement of knee extensors and plantarflexors. The second synergy (peak at ~27% of the stance phase) described the propulsion phase, to which the plantarflexors mainly contributed. The third synergy (peak at ~90% of the stance phase) was associated with the arm swing, when the upper body muscles played an important role. The fourth synergy (peak at ~22% of the swing phase) identified the early swing, showing contributions from upper body muscles, stabilizing muscles of the lower limb and the start of foot dorsiflexors activation. The fifth and last synergy (peak at ~72% of the swing phase) reflected the late swing and the landing preparation, highlighting the relevant contribution of knee flexors, foot dorsiflexors (in the shod condition) and plantarflexors (in the barefoot condition).

The motor primitives of the weight acceptance, propulsion and early swing synergies were significantly dissimilar ($p=0.023$, 0.002 and <0.001, respectively; Figure 23, Table 12). The motor modules exhibited significant differences in the propulsion ($p < 0.001$), arm swing ($p = 0.023$) and late swing ($p < 0.001$) synergies (Figure 23). The muscles responsible for said changes where mainly the upper and lower leg muscles in the propulsion (higher contribution in the shod condition), the trunk muscles in the arm swing (higher contribution in the barefoot condition), the knee flexors and foot plantarflexors in the late swing synergy (higher contribution in the barefoot condition, Figure 23).

The *CoA* of the motor primitives for all the synergies, except from the early swing one, moved significantly in time. The *CoA* values were lower in barefoot running (anticipated activation) for those synergies related completely or partially to the stance phase. For those synergies describing the only swing phase, the *CoA* values were instead bigger in the barefoot compared to the shod condition (Table 13). Further, we found a significant ($p <$

0.001) decrease in the FWHM of the propulsion primitives and an increase ($p < 0.001$) of the arm swing primitives in barefoot compared to shod running (Table 13).

Table 12 Similarities, indicated as $R^2_{S,B}$, between the motor primitives of shod and barefoot running as mean of intraday repetitions. The intraday repeatability values are reported as mean of four trials (two shod and two barefoot). Values ± Type A uncertainty. The *p-values* were calculated by comparing the R^2 between shod and barefoot running and the R^2 for intraday trials.

Motor primitives			
	$R^2_{S,B}$	**$R^2_{S,B}$ intraday**	***p-value***
Weight acceptance	*0.87 ± 0.15*	*0.92 ± 0.11*	*0.023**
Propulsion	*0.91 ± 0.08*	*0.92 ± 0.21*	*0.002**
Arm swing	0.77 ± 0.35	0.82 ± 0.24	0.785
Early swing	*0.82 ± 0.24*	*0.89 ± 0.25*	*<0.001**
Late swing	0.90 ± 0.09	0.88 ± 0.15	0.837

Table 13 Differences between shod and barefoot running in the center of activity (*CoA*) as well as in the relative full width at half maximum (FWHM) of motor primitives. Positive differences ($\Delta_{S,B}>0$) denote bigger values in the barefoot condition, whereas negative differences imply lower values.

	Motor primitives			
	CoA		**FWHM**	
	$\Delta_{S,B}$	***p-value***	**$\Delta_{S,B}$**	***p-value***
Weight acceptance	*-1.3%*	*<0.001**	+3.2%	0.174
Propulsion	*-1.3%*	*<0.001**	*-6.2%*	*<0.001**
Arm swing	*-0.9%*	*0.014**	*+20.2%*	*<0.001**
Early swing	+0.5%	0.271	+1.9%	0.135
Late swing	*+1.2%*	*0.008**	+4.6%	0.554

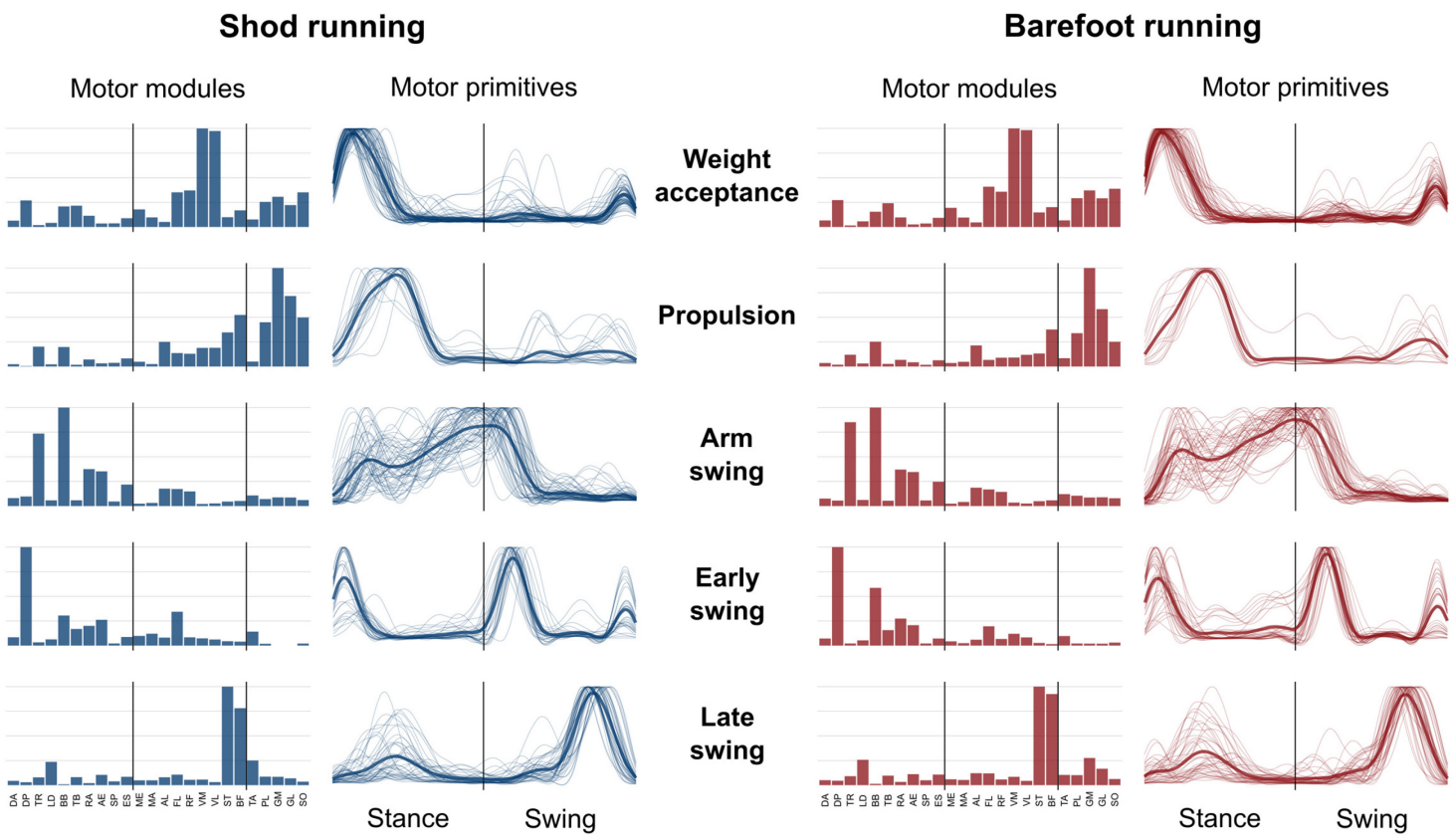

Figure 22 Average motor modules and motor primitives of the five fundamental synergies for shod and barefoot running at the comfort speed. The motor modules are presented on a normalized y-axis base. For the motor primitives, the x-axis full scale represents one gait cycle (stance and swing normalized to the same amount of points and divided by a vertical line) and the y-axis the normalized amplitude. Asterisks denote significant differences between shod and barefoot running.

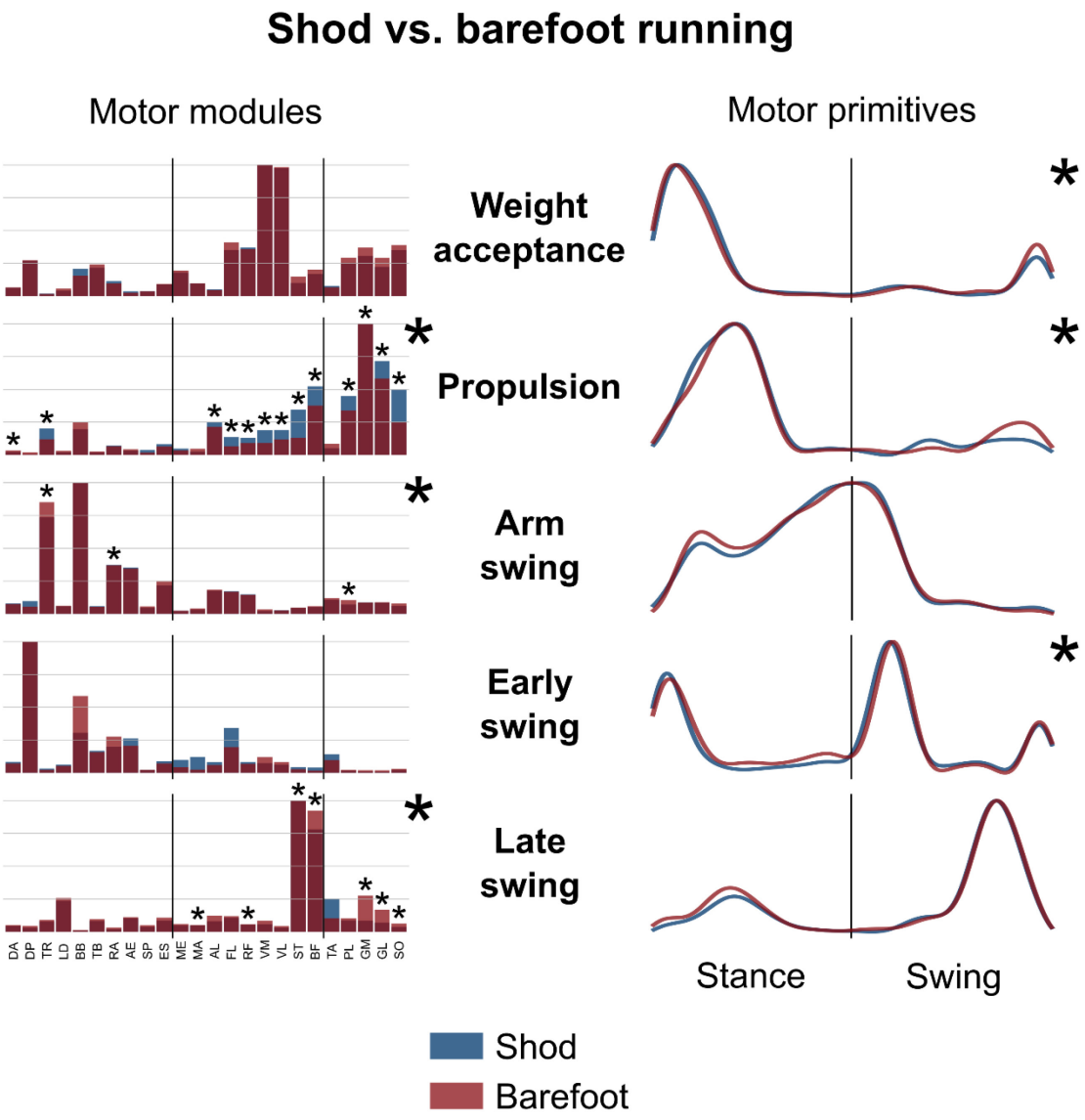

Figure 23 Average motor modules and motor primitives of the five fundamental synergies for shod and barefoot running at the comfort speed. The motor modules are presented on a normalized y-axis base. For the motor primitives, the x-axis full scale represents one gait cycle (stance and swing normalized to the same amount of points and divided by a vertical line) and the y-axis the normalized amplitude. Asterisks denote significant differences between shod and barefoot running.

4.5 Discussion

In this study, we analyzed the modularity of the neuromuscular control of shod and barefoot running. We hypothesized a different modular organization of motion mainly due to the presence or absence of shoes in the two conditions. We found that the motor primitives (or fundamental activation patterns) were generally shifted earlier in time during the stance-related phases and later in the swing-related ones. The motor primitives were found to be significantly wider in the arm swing phase but not in the propulsion, where the basic activation was significantly of shorter duration (peculiarity confirmed by the analysis of the spinal motor output). Moreover, the motor modules (or muscle weightings) demonstrated analogous organization with some significant differences in the propulsion, arm swing and late swing synergies.

The cadence and the strike index significantly increased when changing from shod to barefoot running. Contact times and VGRFs decreased accordingly in the barefoot compared to the shod condition. These results agree with previous studies[204,281] on the comparison of shod and barefoot running. It is well known that the gear ratios of the ankle joint muscles (i.e. the ratio between the ground reaction force and the muscle force moment arms[151]) do not only vary through the running stance phase[151], but also when switching from the shod to the barefoot condition[230]. In the last 20% of the stance phase the gear ratio at the ankle joint is lower during barefoot compared to shod running[230]. Lower gear ratios at the ankle joint decrease the contact time while running[282] and provide an explanation for the shorter contact times found during barefoot running. Further, a lower gear ratio at the ankle joint induces a reduction in the potential of the plantarflexors to generate efficient muscle force due to the force-velocity relationship[151]. In inexperienced runners, this may initiate a dynamic instability in the whole system (including the upper body), requiring stabilization achieved through feedback- as well as predictive-based motor control. We recently found a significant decrease in the dynamic stability of running by switching from shod to barefoot[283]. Moreover, it has been reported that the intrinsic foot muscles show higher absolute activation levels during stance in shod compared to barefoot running[284]. This difference produces an alteration in the longitudinal arch compression during the stance phase, leading to higher recoil capabilities in barefoot running[284]. This increase in the capacity of the foot to store

and return energy is likely an odd feature for the unexperienced barefoot runner and might be another mechanism driving the system to an increased instability.

These very same factors (i.e. different gear ratios, dynamic stability and foot's recoil capabilities) could as well partly explain the differences we found in the duration of the motor primitives. First of all, the reduction in duration of the propulsion-related primitive might be a direct consequence of the lower gear ratios and, possibly, of the increased energy storage and return capabilities when running barefoot. However, this does not explain the increase in the duration of the motor primitives in the arm swing synergy. It has been recently shown that the FWHM of EMG activity undergoes, during gait, a systematic decrease with age in typically developing children[279]. Conversely, very limited age-related changes appear in children affected by cerebral palsy. Moreover, cerebral palsy and typically developing children show a comparable structure of motor modules[279]. Analogously, a widening of the motor primitives can be found in adult patients with cerebellar ataxia and in healthy adults walking on a narrow beam and on slippery ground[106]. This consolidation of the motor output, promoted by learning and impaired by pathology, might reflect the system's need of adding fail-safe robustness to cope with previously unexperienced running conditions (e.g. the absence of footwear).

Concerning motor modules, significant differences were found in the propulsion, arms swing and late swing synergies. The modules of the propulsion phase indicated that upper leg muscles and, most importantly, foot plantarflexors mainly contributed to the inequality. The relative contributions of these muscles were lower in the barefoot condition, indicating a higher specificity of the muscles more important for the propulsion. During arm swing, the TR, RA and PL muscles were found to be significantly responsible for the identified changes. The relative contribution of TR and PL was higher in barefoot compared to shod running, while the contrary emerged for the RA. However, the intrinsic variability of this synergy's patterns is high and the EMG activities low compared to other gait cycle phases. Therefore, small adaptations in the strategy might translate in statistical differences. The changes in FSP are the cause for the alteration of the motor modules of foot dorsiflexors and plantarflexors in the late swing synergy. In agreement with the prediction based on one of our earlier studies[196], 70% of the participants changed FSP without undergoing a specific training intervention when

switching from shod to barefoot running. Most of the times participants automatically switched from RS (shod) to MFS (barefoot). In some cases, participants changed FSP after a few steps, reportedly due to the discomfort of striking the ground with the bare rearfoot. Specifically, 14 out of 20 participants transitioned from RS (shod) to MFS (barefoot). It is well known that the muscles TA and GM and GL play an important role in the final part of the swing phase, just before touchdown[270]. In RS patterns, the TA has the twofold task of dorsiflexing the foot to prepare it for the strike and to control the plantarflexion immediately after the touchdown[268]. In MFS patterns, given the substantial impact loads at contact during running, a preactivation happens right before the strike and the subsequent activation in the early stages of the stance phase[269]. Looking at the late swing synergy, it is evident that the TA contributed more in shod running, an activity that mostly involves a RS pattern. In contrast, the average pattern in barefoot running was a MFS, where the preactivation of GM and GL is predominant. These considerations might as well be extended to overground running, since it has been recently shown that treadmill and overground running share similar motor modules with minimal temporal shifts in the motor primitives[147].

We cannot exclude that habitual barefoot runners might be able to compensate for the differences in the modular organization of muscle activation found in our participants (which were all inexperienced barefoot runners). Although some effects of barefoot running habituation on FSP can be expected, we argue that the main alterations in the motor modules would remain visible in habitual barefoot runners. This mainly because of the predictable changes in the EMG activity[268–270] and, consequently, in the motor modules associated to the kinematic and kinetic alterations induced by a MFS compared to a RS. Concerning motor primitives, however, we suggest that a training intervention focused on the practice of barefoot running might lead to an improvement in the accuracy of motor commands' timing, thus reducing the FWHM of those primitives that here appear wider. Given the characteristics of barefoot running that we discussed above, however, we do not expect that a retraining program would be able to affect consistently the propulsion motor primitive.

With this study using the muscle synergies concept, we bring new insights in the modular organization of shod and barefoot running. Investigating the differences between the synergies, we could confirm that although in both shod and barefoot running five

fundamental synergies are enough to describe the running task, a dissimilarity exists in the modular organization of movement. Moreover, we found an increase in the FWHM of the motor primitives of the arm swing synergy: a possible indication of weak motor learning[279]. These findings suggest a reorganization of the motor output possibly due to the nervous system's effort to cope with the biomechanical specificity of barefoot running. This specificity might be explained by a lower ankle gear ratio[151], different FSP[268–270] and increased instability[283] created by the absence of shoes. The results indicate a possible reorganization of movement when task's complexity either increases or is not well mastered. The required adjustments seem to go in the direction of an improved robustness of motor output guaranteed by longer activation patterns applied on similar muscle modules, showing some adaptability in such a task-specific structure as the muscle synergies.

4.6 Acknowledgements

The authors are grateful to the participants that always showed great commitment and interest during the measurements and to DAAD for the financial support during the publication process.

4.7 Author contributions

Conceptualization: AS, LJ, AE and AA; Methodology: AS, LJ and AA; Investigation: AS, LJ and AE; Formal Analysis: AS and LJ; Writing – Original Draft: AS; Writing – Review & Editing: AS, LJ, AE, VB and AA; Visualization: AS; Supervision: VB and AA.

5 Fourth study – Challenging human locomotion: stability and modular organisation in unsteady conditions

Alessandro Santuz[1,3], Antonis Ekizos[1,3], Nils Eckardt[2], Armin Kibele[2], and Adamantios Arampatzis[1,3]

[1]Department of Training and Movement Sciences, Humboldt-Universität zu Berlin, Berlin, 10115, Germany
[2]Institute for Sports and Sports Science, Universität Kassel, Kassel, 34121, Germany
[3]Berlin School of Movement Science, Humboldt-Universität zu Berlin, Berlin, Germany

This is a post-peer-review, pre-copyedit version of an article published in Scientific Reports, 2018, 8 (2740). The final authenticated version is available online at: https://dx.doi.org/10.1038/s41598-018-21018-4.

5.1 Abstract

The need to move over uneven terrain is a daily challenge. In order to face unexpected perturbations due to changes in the morphology of the terrain, the central nervous system must flexibly modify its control strategies. We analysed the local dynamic stability and the modular organisation of muscle activation (muscle synergies) during walking and running on an even- and an uneven-surface treadmill. We hypothesized a reduced stability during uneven-surface locomotion and a reorganisation of the modular control. We found a decreased stability when switching from even- to uneven-surface locomotion ($p<0.001$ in walking, $p=0.001$ in running). Moreover, we observed a substantial modification of the time-dependent muscle activation patterns (motor primitives) despite a general conservation of the time-independent coefficients (motor modules). The motor primitives were considerably wider in the uneven-surface condition. Specifically, the widening was significant in both the early (+40.5%, $p<0.001$) and late swing (+7.7%, $p=0.040$) phase in walking and in the weight acceptance (+13.6%, $p=0.006$) and propulsion (+6.0%, $p=0.041$) phase in running. This widening highlighted an increased motor output's robustness (i.e. ability to cope with errors)

when dealing with the unexpected perturbations. Our results confirmed the hypothesis that humans adjust their motor control strategies' timing to deal with unsteady locomotion.

5.2 Introduction

Although the mechanisms underlying human movement are not yet understood in their entirety, there is great amount of information on how we walk and run over even, solid surfaces[91,149–151]. Yet, daily-life locomotion is regularly taking place in more complex environments and conditions which imply facing an extraordinary amount of variables and interactions between them. Being able to effectively move over diverse terrain conditions has been integral part of the evolution of the genus *Homo*, not only to hunt and gather, but to escape predators and find mates as well[152,153]. It is well known that, for animals, one fundamental way to optimise locomotion is minimising the energy cost of motion[127,154,155]. However, optimisation of economy does not fully explain the movement strategies employed by humans and other animals[166]. Increased energy expenditure has been often found for walking and running on irregular natural or artificial terrain as swamp, loam, stubble, grass, sand, snow, rubber, mountain trail and uneven-surface treadmill[156–165]. Studies showed that destabilising environments can decrease dynamic stability during walking[170–172]. Further, in presence of perturbations in the mediolateral direction, participants do not modify their velocity as a coping mechanism for stability, but rather alter spatiotemporal gait parameters[170,173]. It therefore seems that to overcome the non-predictability of uneven terrains, humans have to face constant changes in their locomotor patterns, which may affect the priorities of system.

The ability to move involves a flexible integration of peripheral sensory information into the central nervous system (CNS)[285]. Attempts to describe the biomechanics and energetics of balance control following single- and multiple-step perturbations have been made in both humans and animal models[59,106,176–178,108,116,127,163,164,168,174,175]. Though, very little is known about the neuromuscular strategies employed by the CNS to cope with external perturbations induced by continuous variations in terrain morphology. Lowering the energy cost of unsteady locomotor movements might not constitute the highest priority during locomotion and compromises to achieve, for instance, safety or perceived comfort could be of higher

importance[166–169]. In the natural environment, animals must maintain dynamic stability when facing unexpected perturbations[168,179]. Locomotion on an unsteady terrain could challenge dynamic stability and, thus, the existing neuromuscular strategies. Alterations in the dynamic stability, continuously elaborated via sensory feedback information, provide a feed-forward drive to the CNS to implement the needed adjustments. To use a term derived from systems and software engineering, "robustness" defines "the degree to which a system or component can function correctly in the presence of invalid inputs or stressful environmental conditions"[286]. Likewise, a biological system is evolutionally robust when its characteristics can withstand perturbations or uncertainty[287]. In a similar fashion, we define "robustness" as the ability of the CNS to cope with unexpected perturbations or errors of execution. To better understand unsteady movement, the contribution of the neuromuscular circuit to the production and control of robust movement must be taken into account.

One of the most common hypothesis is that the CNS might simplify the production of movement by activating muscles in common patterns called synergies[9,52,241]. This strategy would avoid the separate activation of each muscle by linearly combining a small set of time-dependent commands[52,241]. Synergies can be seen as low dimensional units that via descending or afferent pathways, produce a complex electromyographic (EMG) pattern in muscles[58,242], creating a locomotor drive mediated by a certain amount of supraspinal control[244]. During walking, the same amount of basic activation patterns could be found in patients with spinal cord injury and in healthy participants at different speeds and gravitational loads[272]. Synergies similar to those found in humans at a spinal[272] or muscular level can be observed also in the motor cortex of the primate and cat[135,186]. This suggests a high degree of cooperation within the central nervous system's structure at all levels. We used an unsupervised learning method called non-negative matrix factorisation (NMF)[56] for reducing the high dimensional EMG input into a small number of synergies.

A few studies already used the muscle synergies concept for investigating the modular control of balance following single-step[116] and multiple-step[106] overground perturbations during walking. Oliveira and colleagues[116] showed that the modular organisation is preserved following a single-step perturbation during walking, while temporal activation patterns might differ. In another study[106], Martino and colleagues found a preservation of the

modular organisation after multiple-step overground walking on slippery ground and narrow beam. Contextually, they found an increase in the width (duration) of the muscle activity patterns during perturbed walking[106], indicating a less accurate control. With "less accurate" we mean less precise in relation to the optimum or, in other words, including a higher degree of random errors that might move the system's state more distant from the centre of the stability attractor[287]. The limitations of these studies, however, mainly lie in the small number of steps recorded, since the muscle synergies results might be influenced by the number of steps involved in the analysed data. Previous studies showed that a high number of steps (≥40) is important to increase the accuracy and precision of the muscle synergies analysis[92]. In order to increase the number of recorded steps and to accurately control the locomotion speed, the use of an uneven-surface treadmill is the most intuitive solution. Moreover, the continuously variable and unexpected perturbations would exclude any short-term predictive behaviour of the system and would help in the generalisation of our results. Exposure to continuously induced perturbations would challenge the system and might require alternative motor control strategies to be employed in the neuromuscular level. To support our analysis of the modular organisation of EMG activity, we investigated the spatiotemporal organisation of the alpha-motoneurons pools in the spinal cord[272]. While this method cannot help in characterising the central pattern generator circuitry, it is a useful tool to better understand the organisation of the motor output at a segmental level, rather than in terms of individual muscle activations.

We have limited knowledge on how humans choose from the available control strategies while moving through complex terrain. The purpose of the current study was to improve our understanding of the mechanisms underlying the neuromuscular control of human unsteady locomotion by investigating walking and running on an even- and an uneven-surface treadmill (Supplementary video 1). We hypothesised an increased instability in both walking and running on the uneven-surface treadmill compared to the even one. We expected that humans would respond to the increased instability by employing a different modular organisation of the neuromuscular output when walking and running on the uneven-surface treadmill. In particular, we expected a transfer from an accurate to a more robust neural control in unsteady locomotion.

5.3 Methods

5.3.1 Experimental protocol

Eighteen healthy and young adults were recruited (11 male, 7 female, height 176 ± 7 cm, body mass 71 ± 13 kg, age 24 ± 3 years, means ± s.d.). All the participants were regularly active and did not use orthotic insoles. None had any history of neuromuscular or musculoskeletal impairments, or any head or spine injury at the time of the measurements or in the previous six months. This study was reviewed and approved by the Ethics Committee of the University of Kassel (approval code E05201602). All the participants gave written informed consent for the experimental procedure, in accordance with the Declaration of Helsinki.

The recordings were conducted on two treadmills: a standard one, equipped with an even-surface (ES) belt (Laufergotest, Erich Jäger, Würzburg, Germany) and one equipped with a custom-made uneven-surface (US) belt (Woodway®, Weil am Rhein, Germany, Figure 24). The US treadmill's belt consisted of terrasensa® classic modules (Sensa® by Huebner, Kassel, Germany) aiming to simulate uneven ground conditions. The kinematics data were recorded through a six-camera motion capture system (Oqus 3+, Qualisys AB, Gothenburg, Sweden) operating at 300 Hz. The muscle activity of 13 ipsilateral muscles was recorded using a 16-channel wireless EMG system (myon m320, myon AG, Schwarzenberg, Switzerland), with a frequency of 1000 Hz.

The participants completed two different tasks on the ES and US treadmill, in random order: walking (shod, fixed velocity; 1.1 m/s female, 1.2 m/s male) and running (shod, fixed velocity; 2.0 m/s female, 2.2 m/s male). The difference in velocity between male and female have been chosen after a pilot study in which we estimated the average gender-specific comfortable running and walking velocity on the US treadmill. The participants were instructed to keep looking at a fixed spot in front of them and avoid looking at the treadmill's belt.

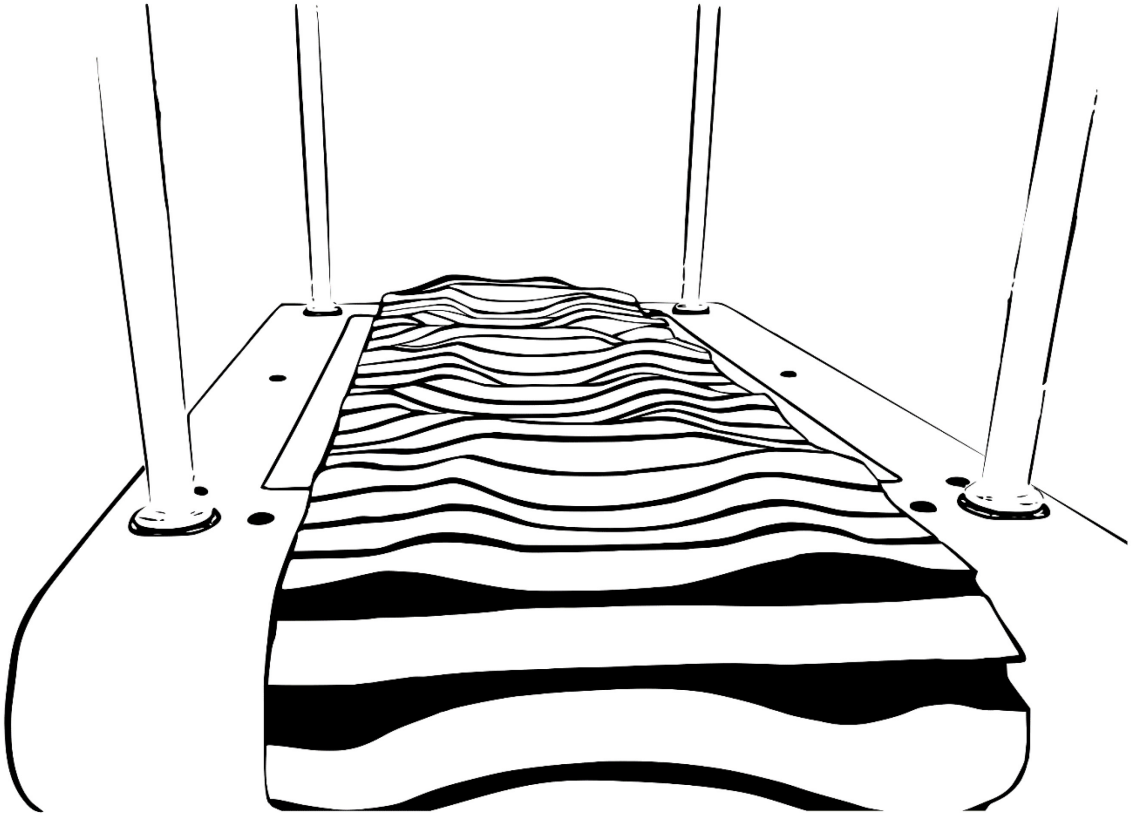

Figure 24 Sketch of the uneven-surface treadmill employed in this study.

5.3.2 Gait cycle assessment

Ten reflective markers were placed bilaterally on the leg, foot and spine. Namely, the greater trochanter, the Achilles tendon insertion on the calcaneus, the dorsal margin of the fifth metatarsal head, the second, seventh and tenth thoracic and the second lumbar *vertebræ* were marked. The gait cycle breakdown was evaluated from kinematic data. We used the information coming from the calcaneus and fifth metatarsal markers. These data were low-pass filtered using a 4th order IIR Butterworth zero-phase filter with cut-off frequency of 50 Hz[288]. For estimating touchdown, we used the modified foot contact algorithm developed by Maiwald *et al.*[288]. For estimating lift-off, we developed the foot acceleration and jerk algorithm. The jerk algorithm searches for the global maximum of the fifth metatarsal vertical acceleration between two consecutive touchdown events to estimate the lift-off (LOe, where the "e" stays for "estimated"). This estimation, however, does not provide an accurate identification of the lift-off and needs some refinement. To get closer to the "real"

lift-off timing, a characteristic minimum in the vertical acceleration (i.e. when the jerk equals zero) of the fifth metatarsal marker is identified in a reasonably small neighbourhood of the LOe. We found [LOe – 90 ms, LOe + 25 ms] for walking and [LOe – 50 ms, LOe + 100 ms] for running to be the sufficiently narrow intervals needed to make the initial lift-off estimation. Both approaches have been validated using force plate data (AMTI BP600, Advanced Mechanical Technology, Inc., Watertown, MA, USA) from 15 participants walking and running overground at six different velocities. True errors were of -8 ± 8 ms (-1.1% ± 1.0% of the stance phase) for touchdown and 12 ± 18 ms (1.6% ± 2.4%) for lift-off for walking and -1 ± 15 ms (-0.5% ± 5.4%) and -16 ± 23 ms (-5.7% ± 8.3%) respectively for running (means ± s.d.).

5.3.3 Local dynamic stability assessment

Participants were allowed for an accommodation period of maximum 60 s[235]. During walking, two trials of 180 s were recorded for all participants, while two trials of 120 s were recorded during the running task. A 4th order IIR Butterworth zero-phase filter with low-pass cut-off frequency of 20 Hz was applied to the computed coordinates of the four spine markers (i.e. the second, seventh and tenth thoracic and the second lumbar *vertebræ*). We adopted the maximum finite-time Lyapunov exponent (MLE) to assess the local dynamic stability of the human system during walking and running. The coordinates of the walking trials were downsampled to 100 Hz to improve computational performance. The vertical coordinates of these markers were then clustered to be used for further analysis and the calculation of the MLE. We calculated the MLE using the vertical coordinate data of the clustered spine markers, which we tested for stationarity[289,290]. Running on a treadmill restricts the movement at the anteroposterior direction due to the participants seeking to match the velocity of the treadmill and similarly, the treadmill width restricts the movement on the mediolateral direction. To avoid dependencies on step frequency, we identified the maximum number of shared steps (i.e. 0.5 of gait cycle) for all trials and participants[291]; 291 and 287 steps for walking and running, respectively. The coordinates of the data segments corresponding to the exact number of steps were then isolated for each trial. Following, the data segments were normalised to a uniform length based on the average number of data for each step, amounting ~16000 in walking and ~33000 data points in running. The high number of steps analysed ensured the reliability of our measurements[292].

State space reconstruction was achieved through delay coordinate embedding[293,294], for each point of the time series and its time-delayed copies as follows:

$$\boldsymbol{S}(t) = [\boldsymbol{z}(t), \boldsymbol{z}(t+\tau), \ldots, \boldsymbol{z}(t+(d-1)\tau)] \qquad (22)$$

with ***S(t)*** being the *d*-dimensional reconstructed state vector, ***z(t)*** the input 1D coordinate series, τ the time delay and *d* the embedding dimension. Time delays were calculated for each time series from the first minimum of the mutual-information curve, based on the Average Mutual Information function[295]. The number of embedding dimensions was extracted through a Global False Nearest Neighbours analysis for each time series, with a threshold of one per thousand data points[296].

Different values of τ and *d* can yield very different state-space reconstructions[289,297,298]. It is therefore suggested that optimised values of τ and *d* are necessary to best represent a dynamical system[283]. In the current study time delays were individually chosen for each participant[283], by first calculating the optimal delay of the four time-series (two trials on the even and two on the uneven surface) and using the averaged value. Time delays were in the range of 17-20 data points (~0.34 of the average step) in the walking task and 36-40 data points (~0.33 of the average step) in the running task.

Following the reconstruction of the times series, the Rosenstein algorithm was used to compute the average exponential rate of divergence by calculating the linear distance of each point's trajectory divergence to its closest trajectory[299,300]. The MLE were then calculated from the slope of the linear fit in the resulting divergence curves from 0 to 1 step. Analysis of the data was performed on MATLAB 2014b (Mathworks Inc., Natick, United States).

5.3.4 Spinal motor output assessment

For each condition, the muscle activity of the following 13 ipsilateral (right side) muscles was recorded: *gluteus medius* (ME), *gluteus maximus* (MA), *tensor fasciæ latæ* (FL), *rectus femoris* (RF), *vastus medialis* (VM), *vastus lateralis* (VL), *semitendinosus* (ST), *biceps femoris* (long head, BF), *tibialis anterior* (TA), *peroneus longus* (PL), *gastrocnemius medialis* (GM), *gastrocnemius lateralis* (GL) and *soleus* (SO). We recorded two trials of 60 s for each condition and analysed the first 50 gait cycles of each acquisition. The EMG

signals were high-pass filtered and then full-wave rectified and low-pass filtered using a 4th order IIR Butterworth zero-phase filter with cut-off frequencies 50 Hz (high-pass) and 20 Hz (low-pass for the linear envelope) using R v3.4.1 (R Found. for Stat. Comp.). The amplitude was normalised to the maximum activation recorded for each participant across all conditions[52,78,262]. Each gait cycle was then time-normalised to 200 points[279], assigning 100 points to the stance and 100 points to the swing phase.

For the spinal motor output characterisation, we mapped the EMG activity onto the estimated rostrocaudal location of alpha-motoneurons (MNs) pools in the segments from the second lumbar vertebra (L2) to the second sacral vertebra (S2) of the spinal cord[272,277,278]. The contribution of each muscle to the total estimated activity of the spinal segments was implemented using the myotomal charts developed by Kendall *et al.*[280]. This method shows the organisation of the efferent MNs network directed to the muscles, assuming a common spinal topography among the investigated participants. The motor output of each spinal segment S_j was estimated using the Equation 23 introduced by La Scaleia *et al.*[278]:

$$S_j = \frac{\sum_{i=1}^{m_j}\left(\frac{k_{ji}}{n_i} \times EMG_i\right)}{\sum_{i=1}^{m_j}\left(\frac{k_{ji}}{n_i}\right)} \times MN_j \qquad (23)$$

where m_j are the muscles innervated by each segment, n_i is the number of spinal levels that innervate the i^{th} muscle, k_{ij} is a weighting coefficient specific to each muscle and spinal segment (e.g. $k_{ij} = 1$ or $k_{ij} = 0.5$ if S_j is a major or minor MN source, respectively) and EMG_i is the normalised recorded EMG, specific for each participant and trial[278,280]. This approach accounts for size differences at each spinal level in every MN_j pool.

5.3.5 Modular organisation assessment

Muscle synergies data were extracted through a custom script[75] (R v3.4.1, R Found. for Stat. Comp.) using the classical Gaussian NMF algorithm[56,75,77] from the first 50 gait cycles of each acquisition[92]. EMG data were pre-processed using the same filtering conditions reported in the previous paragraph and each gait cycle was time-normalised to 200 points[279], assigning 100 points to the stance and 100 points to the swing phase[77]. The $m = 13$ time-dependent muscle activity vectors were grouped in an $m \times n$ matrix V (n equal to 10000),

factorised such that $V \approx V_R = WH$. V_R represents the new reconstructed matrix, which approximates the original matrix V. The motor primitives[75,90] matrix H contained the time-dependent coefficients of the factorisation with dimensions $r \times n$, where r represents the minimum number of synergies necessary to reconstruct the original signals (V). The motor modules[75,253] matrix W with dimensions $m \times r$, contained the time-invariant muscle weightings. H and W described the synergies necessary to accomplish a movement. The update rules for H and W are presented in Equation 24 and Equation 25.

$$\begin{cases} H_{i+1} = H_i \dfrac{W_i^T V}{W_i^T W_i H_i} & (24) \\ W_{i+1} = W_i \dfrac{V(H_{i+1})^T}{W_i H_{i+1}(H_{i+1})^T} & (25) \end{cases}$$

The limit of convergence was reached when a change in the calculated R^2 between V and V_R was smaller than the 0.01% in the last 20 iterations[75,79]. This was done for a number of synergies successively increased from 1 to 10. The computation was repeated 10 times for each synergy, each time creating new randomised initial matrices H and W, in order to avoid local minima[75,80]. The solution with the highest R^2 was then selected for each of the 10 synergies.

To choose the minimum number of synergies required to represent the original signals, the curve of R^2 values versus synergies was fitted using a simple linear regression model, using all ten synergies. The mean squared error[75] was then repeatedly calculated, each time removing the lower synergy point, until only two points were left or until the mean squared error fell below 10^{-5}. The aforementioned procedure allowed us to extract fundamental and combined synergies from the raw EMG data. A fundamental synergy can be defined as an activation pattern whose motor primitive shows a single peak of activation[75,77]. When two or more fundamental synergies are blended into one, a combined synergy appears. Combined synergies were excluded from the analysis. An example of combined synergies is reported in Figure 25.

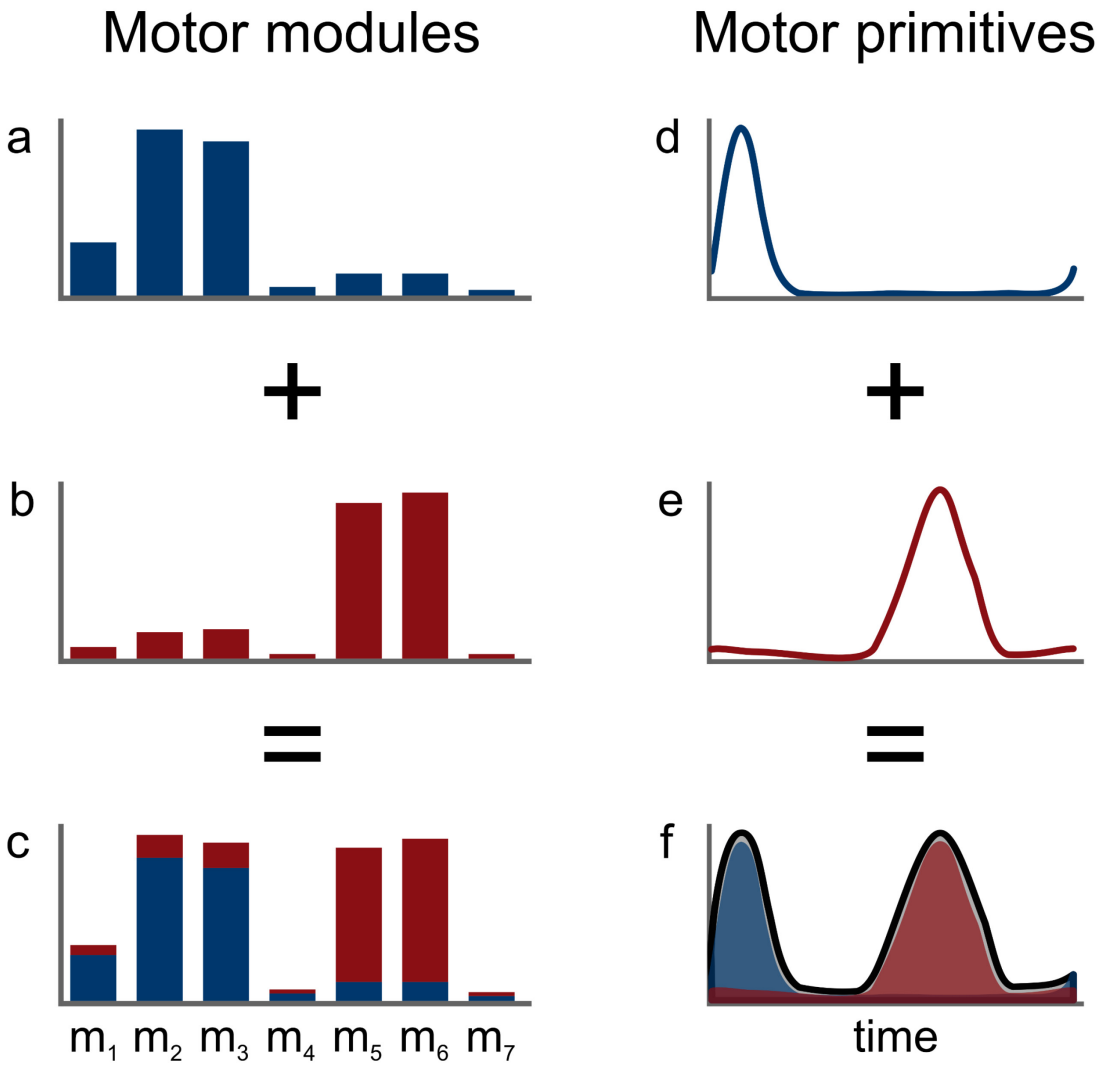

Figure 25 Example of two fundamental synergies combined into one. The histograms a and b represent the two fundamental sets of motor modules for seven muscles. The curves d and e show the two respective primitives, with arbitrary x- and y-axis units. The combined motor modules and primitives are presented in panels c and f, respectively.

5.3.6 Metrics for comparison of curves

We evaluated the centre of activity (*CoA*) and full width at half maximum (FWHM) for the resulting curves of the extracted spinal maps and motor primitives (matrix H) in both conditions and types of locomotion. The *CoA* was defined as the angle of the vector (in polar coordinates) that points to the centre of mass of that circular distribution[77,279]. The polar direction represented the gait cycle's phase, with angle $0 \leq \theta_t \leq 2\pi$. The following equations define the *CoA*:

$$\begin{cases} A = \sum_{t=1}^{p} (\cos \theta_t \times P_t) & (26) \\ B = \sum_{t=1}^{p} (\sin \theta_t \times P_t) & (27) \end{cases}$$

$$CoA = \arctan(B/A) \qquad (28)$$

where p is the number of points of each gait cycle (p = 200) and P is the activation vector. The FWHM was calculated as the number of points exceeding each gait cycle's half maximum, after subtracting the gait cycle's minimum[77,279]. The first 50 gait cycles of each acquisition were selected for analysis. The *CoA* and FWHM were analysed step-by-step and then averaged for stance and swing distinctively for spinal maps and over the whole gait cycle for the motor primitives.

5.3.7 Statistics

To evaluate the differences in MLE and gait parameters between ES and US locomotion, we used a one-way ANOVA with repeated measures or non-parametric Friedman test in case the normality assumptions on the residuals were not satisfied. To investigate *CoA* and FWHM of the spinal motor output, we employed a two-way ANOVA with repeated measures, followed by a Tukey *post-hoc* analysis with false discovery rate *p-value* adjustment (the independent variables being the condition, i.e. ES or US, and the spinal segment). The same statistics was used to assess differences between the motor modules, using the muscles and the conditions (ES or US) as independent variables. To evaluate the similarities between the fundamental

motor primitives of the ES and US conditions, we used the coefficient of determination R^2. The analysis was conducted as follows: first, we calculated the similarity between the pairs of trials recorded during ES and US locomotion (i.e. the repeatability level when comparing two trials of the same condition)[75,77]. Then, we statistically compared the similarity between trials of different conditions (ES and US locomotion) with the repeatability values. To do so, we used a one-way ANOVA with repeated measures or non-parametric Friedman test in case the normality assumptions on the residuals were not satisfied. Type A uncertainty was expressed as $u_A = s/\sqrt{n}$. All the significance levels were set to $\alpha = 0.05$ and the statistical analyses were conducted using R v3.4.1 (R Found. for Stat. Comp.).

5.3.8 Data availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

5.4 Results

5.4.1 Gait parameters

The contact times in walking did not differ significantly ($p = 0.539$) between ES (674 ± 37 ms) and US (668 ± 49 ms). In running, there was a significant decrease ($p < 0.001$) in contact time when switching from ES (353 ± 50 ms) to US (324 ± 41 ms). Similarly, cadence did not change in walking ($p = 0.589$, 109 ± 5 and 110 ± 7 steps/min for ES and US, respectively), while in running it increased significantly ($p < 0.001$) when transitioning from ES (155 ± 7 steps/min) to US (160 ± 10 steps/min, means ± s.d.).

5.4.2 Local dynamic stability

The MLE were significantly higher in the US condition in both walking (ES: 1.557 ± 0.087, US: 1.655 ± 0.119, $p < 0.001$) and running (ES: 1.936 ± 0.072, US: 2.031 ± 0.085, $p = 0.001$, means ± s.d.), evidencing an increased instability in the US (Figure 26).

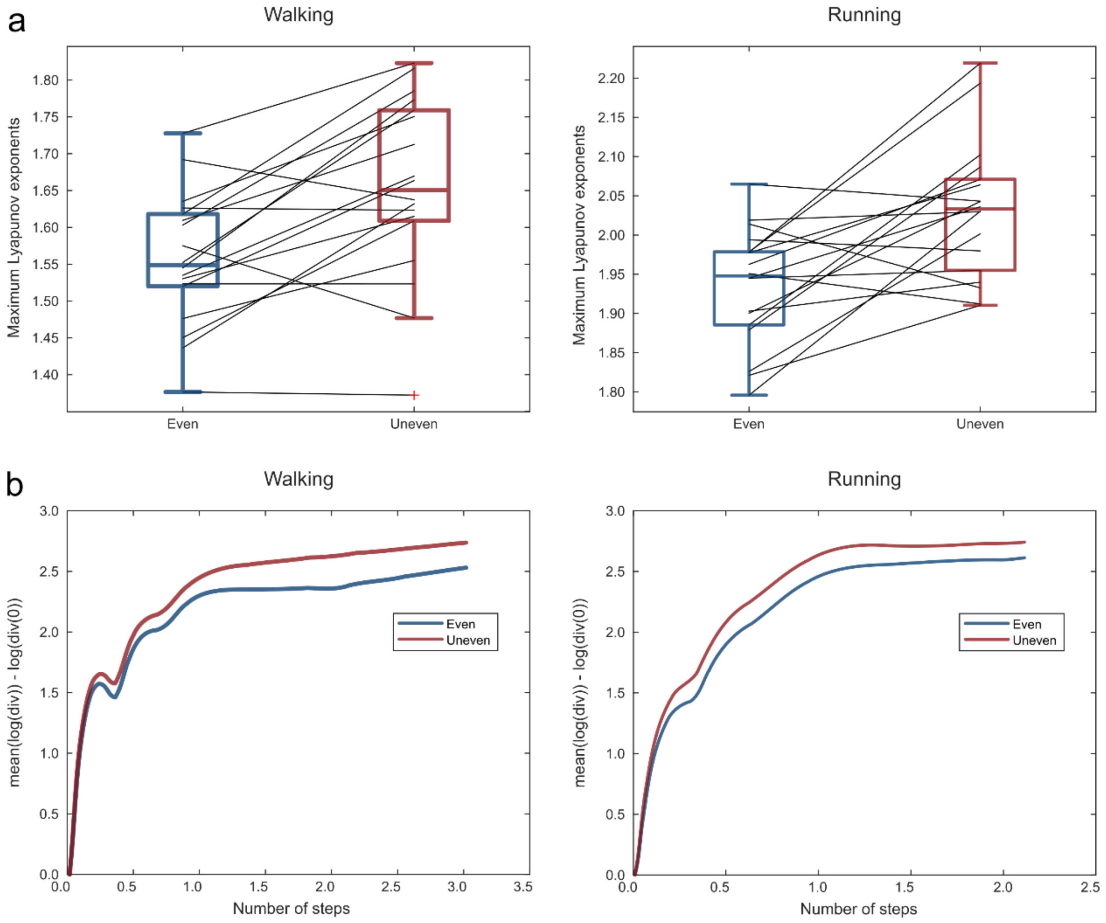

Figure 26 Boxplots depicting the maximum Lyapunov exponent values for even and uneven surface in both walking and running with the black lines indicating individual changes between conditions (a). Averaged over all participants mean logarithmic divergences of the trajectories between even and uneven surface in walking and running. A faster divergence indicates worse dynamic stability (b).

5.4.3 Spinal motor output

Figure 27 depicts the average spinal motor output for ES and US walking and running. The two-way ANOVA identified statistically significant differences ($p < 0.001$, Table 14) in the *CoA* of the mapped EMG activities when comparing ES and US locomotion only for the stance phase of the walking cycle. The *post-hoc* analysis showed a significant contribution ($p < 0.001$) of the S3 spinal segment. The FWHM of the mapped EMG activities was significantly different between ES and US conditions for the stance as well as the swing phase of walking ($p = 0.002$ and $p=0.001$ respectively, Table 15). For running, the FWHM

was significantly different ($p < 0.001$) only in the swing phase (Table 15). The *post-hoc* analysis evidenced significantly greater FWHM values ($p < 0.001$) in the US condition in the S3 spinal segment for walking (stance phase). For running, the segments L4, L5, S1, S2 in the swing phase were involved instead (Table 15).

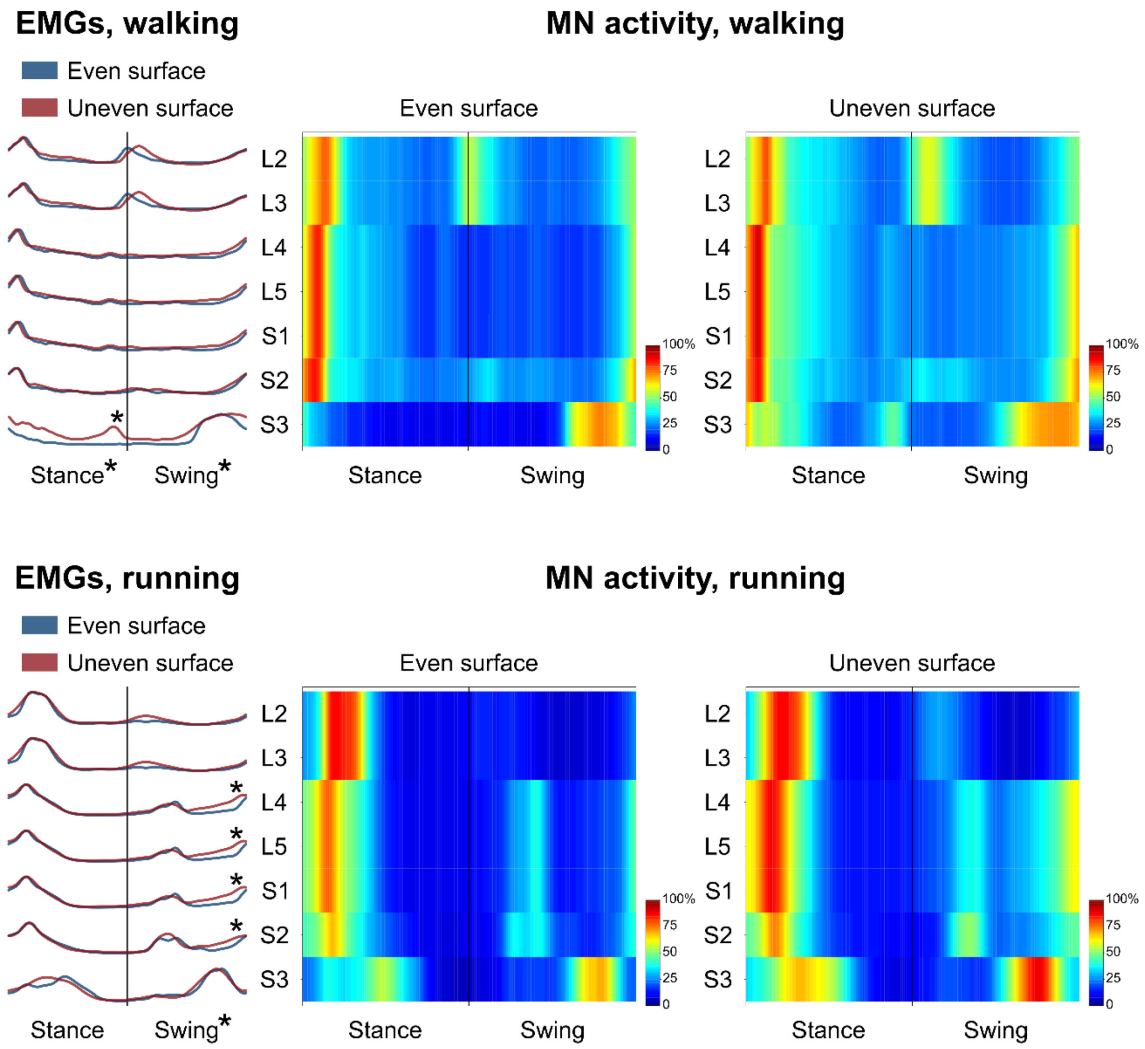

Figure 27 The average spatiotemporal spinal motor outputs are presented for even and uneven surface walking and running, normalised in amplitude to the maximum of each segment. These curves have been obtained by mapping each of the 13 muscle activations onto the relevant spinal segment (lumbar from L2 to L5 and sacral from S1 and S3). Asterisks denote significant differences in the full width at half maximum of the mapped EMGs between even and uneven surface locomotion. The two level plots show the average alpha-motoneurons activity for each condition, giving additional information about the absolute activation level (normalisation to the maximum of each condition). The stance and swing phases have been temporally normalised to the same amount of data points (100 each). Values are the means across all subjects and all trials.

Table 14 Differences between even and uneven surface in the centre of activity (*CoA*) of the EMG activities mapped onto the estimated rostrocaudal location of the spinal cord (segments L2 to S3). Positive differences ($\Delta_{E,U}>0$) denote bigger values in the uneven surface condition, whereas negative differences imply lower values.

	Walking			**Running**	
	Stance ***p*<0.001***		**Swing** ***p*=0.110**	**Stance** ***p*=0.526**	**Swing** ***p*=0.648**
Segment	**$\Delta_{E,U}$**	***p-value***	**$\Delta_{E,U}$**	**$\Delta_{E,U}$**	**$\Delta_{E,U}$**
L2	+1.4%	0.783	-6.8%	+0.7%	-7.1%
L3	+1.4%	0.783	-6.8%	+0.7%	-7.1%
L4	+2.9%	0.416	-0.8%	+0.0%	+4.8%
L5	+2.9%	0.416	-0.8%	+0.0%	+4.8%
S1	+2.9%	0.416	-0.8%	+0.0%	+4.8%
S2	+4.0%	0.285	-0.5%	-0.7%	+3.9%
S3	*+13.7%*	*<0.001**	-3.6%	-3.6%	+0.8%

Table 15 Differences between even and uneven surface in the relative full width at half maximum (FWHM) of the EMG activities mapped onto the estimated rostrocaudal location of the spinal cord (segments L2 to S3). Positive differences ($\Delta_{E,U}>0$) denote bigger values in the uneven surface condition, whereas negative differences imply lower values.

	Walking				**Running**		
	Stance ***p*=0.002***		**Swing** ***p*=0.011***		**Stance** ***p*=0.119**	**Swing** ***p*<0.001***	
Segment	**$\Delta_{E,U}$**	***p-value***	**$\Delta_{E,U}$**	***p-value***	**$\Delta_{E,U}$**	**$\Delta_{E,U}$**	***p-value***
L2	-0.6%	0.610	+2.0%	0.279	+1.2%	+0.1%	0.070
L3	-0.6%	0.610	+2.0%	0.279	+1.2%	+2.7%	0.070
L4	+1.2%	0.252	+1.8%	0.315	+0.8%	*+2.7%*	*<0.001**
L5	+1.2%	0.252	+1.8%	0.315	+0.8%	*+7.3%*	*<0.001**
S1	+1.2%	0.252	+1.8%	0.315	+0.8%	*+7.3%*	*<0.001**
S2	+0.1%	0.942	-1.0%	0.552	-1.0%	*+4.4%*	*0.003**
S3	*+5.9%*	*<0.001**	+2.5%	0.133	+3.4%	+2.0%	0.172

5.4.4 Modular organisation

The minimum number of synergies necessary to sufficiently describe the measured EMG-activity during walking and running was not significantly different between ES and US in either walking (3.8 ± 0.6 for ES and 3.9 ± 0.5 for US, p = 0.665) or running (3.5 ± 0.5 for ES and 3.5 ± 0.7 for US, p = 0.743, means ± s.d.). In both locomotor activities, four fundamental activation patterns could be identified (Figure 28). The four fundamental synergies extracted during ES and US walking and running were associated with temporally different phases of the gait cycle. The first synergy (peak at ~12% and ~21% of the stance phase for walking and running, respectively) functionally referred to the body weight acceptance, with a major involvement of knee extensors and glutei. The second synergy (peak at ~70% and ~46% of the stance phase for walking and running, respectively) described the propulsion phase, to which the plantarflexors mainly contributed. The third synergy (peak at ~8% and ~30% of the swing phase for walking and running, respectively) identified the early swing, showing the involvement of foot dorsiflexors. The fourth and last synergy (peak at ~75% and ~80% of the swing phase for walking and running, respectively) reflected the late swing and the landing preparation, highlighting the relevant influence of knee flexors and foot dorsiflexors.

In walking, the similarities between the ES and US motor primitives of the propulsion and early swing synergies were significantly lower than the intraday repeatability threshold (Figure 28, Table 16). In running, the motor primitives of all synergies were found to be significantly different between ES and US locomotion (Figure 28, Table 16). The motor modules of ES and US locomotion exhibited significant differences in the weight acceptance synergy for walking (p = 0.001) and in the late swing synergy for running (p = 0.046, Figure 28).

The *CoA* of the motor primitives for the propulsion, early swing and late swing synergies moved significantly earlier in time ($p<0.05$) in US compared to ES locomotion (Table 17, the only exception being the late swing phase in running). Further, we found an increase of the FWHM in the US compared to ES condition of the primitives related to early and late swing in walking and to the weight acceptance and propulsion in running (Table 17).

Table 16 Motor primitives' similarities, indicated as $R^2_{E,U}$, between even and uneven surface walking and running as mean of intraday repetitions. The intraday repeatability values are reported as mean of four trials (two on the even and two on the uneven surface). Means ± Type A uncertainty. The *p-values* were calculated by comparing the R^2 between even and uneven and the R^2 for intraday trials.

		Motor primitives		
		$R^2_{E,U}$	$R^2_{E,U}$ intraday	*p-value*
Weight acceptance	Walking	0.78 ± 0.20	0.89 ± 0.09	0.055
	Running	*0.73 ± 0.27*	*0.94 ± 0.09*	*0.001**
Propulsion	Walking	*0.73 ± 0.20*	*0.89 ± 0.10*	*0.004**
	Running	*0.70 ± 0.21*	*0.95 ± 0.03*	*<0.001**
Early swing	Walking	*0.05 ± 0.83*	*0.79 ± 0.13*	*<0.001**
	Running	*-0.34 ± 0.76*	*0.74 ± 0.65*	*<0.001**
Late swing	Walking	0.72 ± 0.18	0.85 ± 0.17	0.059
	Running	*0.73 ± 0.16*	*0.95 ± 0.08*	*<0.001**

Table 17 Differences between even and uneven surface walking and running in the centre of activity (*CoA*) as well as in the relative full width at half maximum (FWHM) of motor primitives. Positive differences ($\Delta_{E,U}>0$) denote bigger values in the uneven surface condition, whereas negative differences imply lower values.

		Motor primitives			
		CoA		**FWHM**	
		$\Delta_{E,U}$	**p-value**	**$\Delta_{E,U}$**	***p-value***
Weight acceptance	Walking	-0.6%	0.478	+4.0%	0.398
	Running	-1.1%	0.112	*+13.6%*	*0.006**
Propulsion	Walking	*-2.5%*	*0.002**	+8.2%	0.180
	Running	*-3.8%*	*<0.001**	*+6.0%*	*0.041**
Early swing	Walking	*-6.7%*	*0.002**	*+40.5%*	*<0.001**
	Running	*-14.1%*	*<0.001**	+10.0%	0.202
Late swing	Walking	*-2.7%*	*0.001**	*+7.7%*	*0.040**
	Running	*+1.6%*	*0.008**	+3.0%	0.212

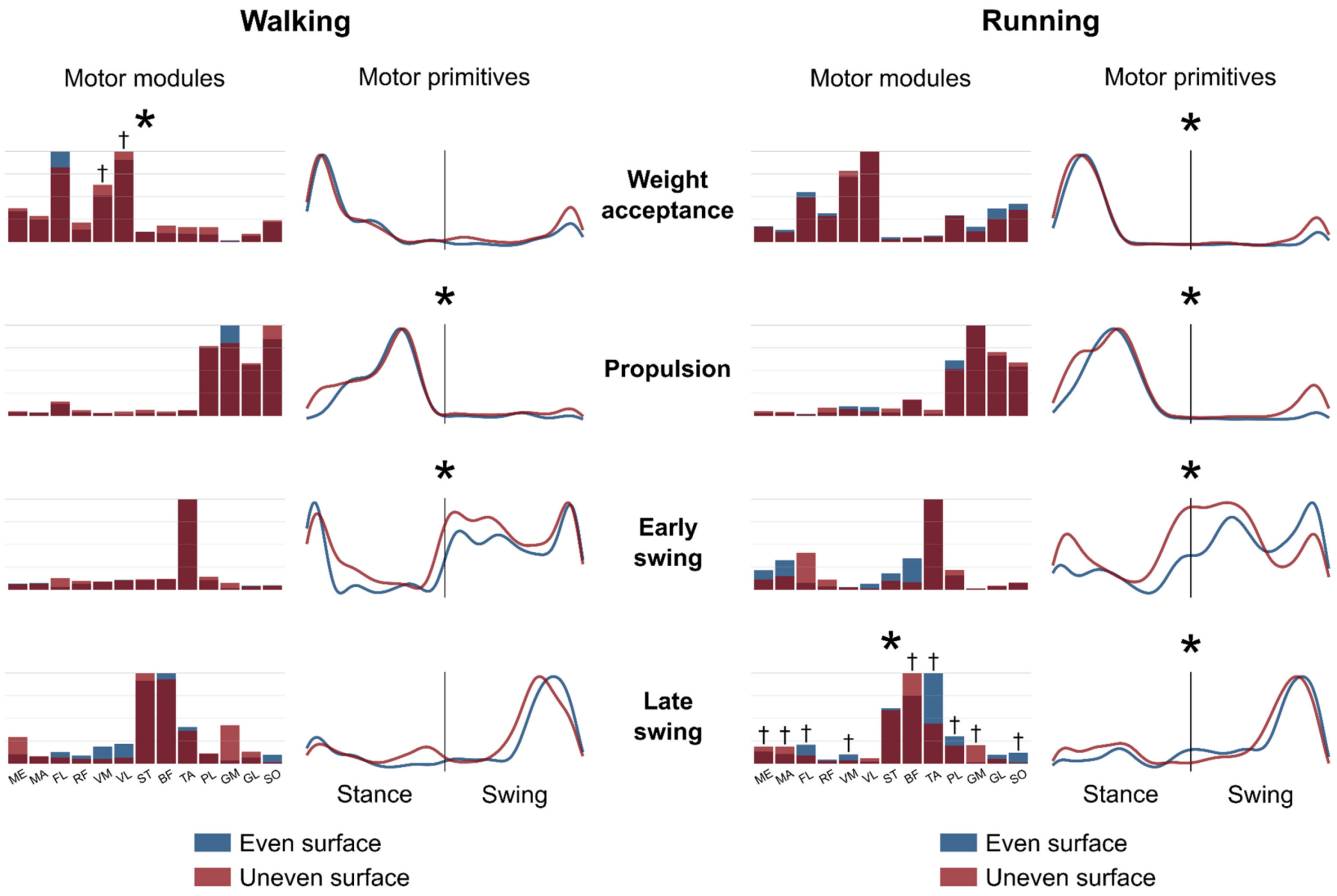

Figure 28 Average motor modules and motor primitives of the four fundamental synergies for walking and running on even and uneven surface. The motor modules are presented on a normalised y-axis base. For the motor primitives, the x-axis full scale represents one gait cycle (stance and swing normalised to the same amount of points and divided by a vertical line) and the y-axis the normalised amplitude. Asterisks denote significant differences between even and uneven surface locomotion. Daggers denote results of the *post-hoc* analysis.

5.5 Discussion

The current study examined the neuromuscular control of normal and perturbed movement during walking and running. We hypothesized a decrease in the dynamic stability and a transfer from an accurate (i.e. mature, functionally fine-tuned) to a more robust (i.e. able to cope with unexpected errors) motor control during US locomotion. We found higher values in the MLE (i.e. higher instability) and a widening in the motor primitives (i.e. alterations in the temporal structure of motor control) and spinal motor output, which evidenced an increased robustness[106,279] of the system during US locomotion. The findings confirmed our hypotheses and demonstrated the use of a consistent set of neural control elements during perturbed steady state locomotion, but with modifications of the basic activation patterns. This indicated a transition from an accurate to a more robust movement control in the presence of continuously variable perturbations.

The MLE quantifies how the system responds to small internal perturbations[291], revealing the ability of the system to maintain stability and detects neuromuscular control errors in achieving it[300,301]. Increased MLE correspond to a more chaotic and unstable dynamical system[289,297]. In our study the MLE increased ~6% in walking and ~5% in running on the US providing evidence for a clear reduction of stability during perturbed locomotion. Previous studies in walking found an increase of 9% in the MLE in patients with focal cerebellar lesion[302] and 21% increase in patients with moderate neurological gait disorders[303], while in running an increase of 2% in the MLE was found during the transition from shod to barefoot[283]. The introduced perturbations, unpredictable and continuously variable in amplitude, interfered with the normal locomotor patterns affecting the neuromuscular control. Based on our findings, the US locomotion resulted in a decreased dynamic stability and likely drove the system to adjust the motor output for robustness. The wider shape of the motor primitives during US locomotion was coupled with a temporal shift in the *CoA*. A similar tendency could be observed as well in the estimated spinal motor output, supporting the idea of a temporal widening also in terms of segmental organisation. Specifically, the spinal motor output in walking was significantly wider especially in the stance phase (but in the swing phase as well), mainly due to the innervations of the muscle *biceps femoris* (spinal segment S3). During running, the width of the spinal motor output

was significantly different only in the swing phase, but for almost all segments and, most importantly, due to the contributions of all muscles' innervations. This result fits with the *post-hoc* analysis conducted on the motor modules of this locomotion type. However, both locomotion conditions showed a similar modular organisation, since four synergies were sufficient to describe the motor task in both ES and US locomotion. These findings provide evidence that the central nervous system uses a consistent set of neural control elements with a flexible temporal recruitment to create safe locomotion in the presence of continuous perturbations during walking and running. The kind of perturbation induced by the US treadmill used in this study was uninterrupted over the acquisition time, thus creating a new perturbation at each step. Since the participants were asked not to look at their feet, we can exclude any proper predictive behaviour (i.e. experience-based prediction of expected perturbations[178]) in all conditions. Nevertheless, we can expect a certain level of anticipation utilised to cope with the potential perturbations. However, they expected continuously variable perturbations and therefore might have been able to create anticipatory muscle activation patterns driven by knowledge and prior experience with the potential perturbation. The main alterations in the modular organisation of the less stable (US) locomotion occurred in the basic activation patterns (motor primitives) rather than in the number of muscle synergies or the structure of the motor modules. Muscle synergies are coordinated patterns of muscle activity that aim to create functional motor output from the interplay of spinal and supraspinal interactions with the environment[52,58]. Synergies might be expressed via motor circuits in the cortex, brainstem and spinal cord[55,135,244]. There is indirect evidence that intrinsic networks of spinal interneurons might be involved in rhythm generation, left-right alternation and flexor-extensor interaction[2,7,45,46,59,180,183,190]. Proofs that a finer, time-dependent tuning of novel or learned elementary spinal commands might be of a supraspinal nature, have been found in the cat[185,186] and monkey[135,187,188]. This suggests that the descending commands (i.e. motor primitives) from the brainstem and motor cortex modulate spinal motor modules[58,59]. Therefore, the widening of the motor primitives indicates a relevant contribution of supraspinal structures in the control of perturbed locomotion.

Previous studies found shorter, faster and wider steps in response to destabilising environments during walking[170,304]. Shorter times to plan and execute movements decrease

their accuracy[173,305]. In our study, similar findings in contact times and cadence (i.e. shorter contact times and higher cadence) were present only during running, possibly due to the fact that the perturbations were continuously induced and affected all planes of motion. This partial inconsistency could be further explained by the magnitude of the induced perturbations, which in our study was not quantifiable and possibly too low to cause modifications in the spatiotemporal parameters of walking. However, when analysing the step-to-step variability, we could find a tenfold increase in the contact times variance in walking ($p<0.001$). To a minor extent, but still significantly ($p=0.001$), this variability was also present in running, where variances increased 2.8 times. However, only in running we could demonstrate a decrease in the average contact times. This displays that the choice of a proper timing of execution is of crucial importance for managing external perturbations. When the phasing of events is less predictable (i.e. stereotyped to an extent that can be managed by a standard set of anticipatory spatiotemporal commands), a loss in accuracy can be expected. Broader basic activation patterns and EMG profiles have already been associated with inaccuracy and variability in motor control as well as with higher metabolic cost in different gait conditions[106,163,164,279]. Therefore, the system maintains successful locomotion by making up for a decrease in accuracy with an increase in robustness, which is reflected in the widening of the motor output.

During development, the locomotor activity undergoes adaptations which are linked to a functional reorganisation of the motor output[279]. As recently reported from Cappellini *et al.*, typically developing children show, during walking, a gradual reduction in the FWHM of motor primitives associated with maturation (i.e. an improvement in accuracy) [279]. Conversely, cerebral palsy children show broader motor primitives compared to typically developing children at the same age[279]. Moreover, in children affected by cerebral palsy the narrowing of the motor primitives with age is lacking, despite a comparable structure of motor modules[279]. Analogously, a widening of the motor primitives can be found in adult patients with cerebellar ataxia and in healthy adults walking on a narrow beam and on slippery ground[106]. The prolongation of the basic activation patterns might reflect the system's need of adding robustness to maintain functionality and overcome continuous perturbations. If we think of motor primitives' widening as a modification of the states of the system, we can assume that functionality is maintained, but the strategy to achieve it has

slightly changed. In fact, it is commonly accepted that a system can achieve robustness by either returning to its current attractor or by moving to a new attractor which is good enough to maintain the system's functions[287]. When a modification in the system's state happens and the system is able to return to its original attractor, there is a so-called "robust adaptation"[287]. Furthermore, it is also true that, in case of a transition to a new attractor, the switch must preserve enough robustness in order to allow the system for consistent and adequate responses to perturbations[287]. This is a crucial feature of robust adaptation, since it allows the system to maintain specific functionalities with different, flexibly selected modes of operations[287]. Under a control systems' perspective, it is known that fuzzy control systems decrease robustness in order to increase optimality and performance[306]. Looking at the widening of the motor primitives in the US condition, it is in fact possible to note that two chronologically adjacent primitives (e.g. weight acceptance and propulsion synergies during running or early swing and late swing during walking) are overlapping more than in ES locomotion. This increases the fuzziness of the temporal boundaries, thus creating a "buffer" of motor control that allows for shifting more easily from one synergy (or gait phase) to the other, increasing the robustness and contextually decreasing optimality and performance[306].

In our study, the widening found in the spinal maps and motor primitives was associated to the less practiced and more unstable of the two tasks (i.e. US walking and running), indicating a less refined timing and duration of the motor output. The requirement to cope with continuously variable perturbations in order to maintain dynamic stability, challenges the neural system's control of locomotion. As such, it may reduce the need for accurate neural control prioritising the search for robustness. The widening of the motor primitives in US walking and running was common among all participants. Concerning motor modules, we could confirm a significant alteration only for those related to the weight acceptance synergy in walking and to late swing in running. This provides an indication of a certain degree of conservation in the basic modular structure. Furthermore, our data confirmed that it is possible to describe with the same number of synergies not only both walking and running[91], but ES and US locomotion as well. This observation provides additional support to the idea that the CNS may be able to modulate existing synergies in order to face different locomotion conditions.

During locomotion, the muscle activation patterns are characterised by a flexible modulation dependent on several external and internal factors. It has been shown by Akay *et al.* that the absence of proprioceptive sensory feedback from muscle spindles and Golgi tendon organs deteriorates the coordination of walking and swimming in mice[190]. A flexible modulation of the motor output has been found in humans undergoing changes in the mechanical demands of walking in both *in vivo*[272] and *in silico* models[120]. Biewener and Daley proposed that mechanical effects are likely to be the predominant motor controllers at high locomotion velocities (e.g. running), since feedback delays might be destabilising[127]. Yet, at lower gait velocities like in walking, proprioceptive sensory feedback might provide superior contribution to stability control[127], a concept that have been confirmed by scale studies on the sensorimotor responsiveness in the giraffe[307]. As mentioned above, we could not exclude some anticipatory behaviour in our participants to cope with the ground changes when walking or running on the US treadmill. The widening of the weight acceptance and propulsion primitives (i.e. alterations directly after touch down) in US running indicates an amplified anticipatory adjustment compared to walking (widening in the early and late swing primitives). On the one hand, appropriate anticipatory adjustments during perturbed locomotion at high velocity could have forced the system to rely more on the intrinsic mechanical response than on the sensory feedback during the stance phase. On the other hand, the widening in walking was present only in the two swing phases (i.e. when the system stands on one leg). This suggests an increase in the robustness of the swinging leg's neuromuscular control when the contralateral limb is reinforcing the proprioceptive sensory feedback by exchanging forces with the ground. Thus, our findings would support the speed-dependent strategies for controlling locomotion hypothesized by Biewener and Daley.

In conclusion, the findings of the current study provide evidence that humans adjust their motor control strategies when walking or running over uneven terrain. We found that the changes in terrain morphology decreased the dynamic stability of the system, resulting in a temporal rearrangement (widening) of the motor primitives' shape, despite a general preservation of the motor modules' structure. The widening indicates an increase in the system's robustness to deal with the induced perturbations. These observations suggest that supraspinal processes might be largely involved in the control of unsteady locomotion, with possible differences in the utilisation of proprioceptive sensory feedback between walking and running.

6 Conclusions and perspectives

With this thesis I endeavoured to provide an insight into the neuromuscular control of perturbed human locomotion. The first task was to conduct a reliability analysis of a computational method for assessing spatiotemporal gait parameters using plantar pressure distribution data (First study – A Pressure Plate-Based Method for the Automatic Assessment of Foot Strike Patterns During Running, page 30). Then, following a similar approach, the influence of EMG data filtering parameters and NMF algorithm's choice on the output of a muscle synergies analysis was investigated (Second study – On the Methodological Implications of Extracting Muscle Synergies from Human Locomotion, page 48). Subsequently, the consolidated methods were applied to the analysis of the modular organisation of perturbed locomotion. Initially, muscle synergies were extracted from the EMG activities created during shod and barefoot running, considering the absence of shoes as a form of mild perturbation of motion in unexperienced barefoot runners (Third study – The Influence of Footwear on the Modular Organization of Running, page 72). Subsequently, synergies were extracted from walking and running over a custom uneven-surface treadmill, which induced uninterrupted and unpredictable disturbances at each step (Fourth study – Challenging human locomotion: stability and modular organisation in unsteady conditions, page 94). In the following paragraphs, the practical implications of the studies I completed during my PhD will be discussed, also addressing their limitations. Finally, I will consider the possible implications of these findings and elaborate on the scientific perspectives that emerged from this work.

6.1 Foot strike patterns during running

There is great debate around the possible benefits of switching from a rearfoot to a mid- or forefoot strike pattern[308,309]. Indeed, a change in the FSP is not only a matter of modifying the ankle joint angle. The whole body weight distribution changes, being the landing point in a rearfoot strike pattern farther from the body's centre of mass than it is in a mid- or forefoot strike pattern[309]. Moreover, rates of force development are different, specifically lower in the mid- or forefoot strike pattern compared to the rearfoot[196,204,308,309]. Also, the peak force values are often higher in rearfoot strikers[308]. Additionally, when rearfoot strikers

switch to a mid- or forefoot strike, usually the cadence (number of steps per minute) increases together with the flight time, while contact times decrease[77,308,309]. Of course, given that the majority of people automatically switch from a rearfoot to a mid- or forefoot strike pattern when taking off their shoes, barefoot running or running in minimalist shoes are a good way to acutely modify the FSP[283,310–312]. Yet, a recent review by Jo Hamill and Allison Gruber[308] underlined the lack of scientific proof about the potential improvements in running economy or injury risk's reduction of mid-/forefoot strike against rearfoot strike patterns. Irene Davis, Daniel Lieberman and many other scientists are nevertheless pushing their research as advocates of barefoot or minimalist running and the question posed at the beginning of this paragraph is still far from being satisfactorily answered[204,309].

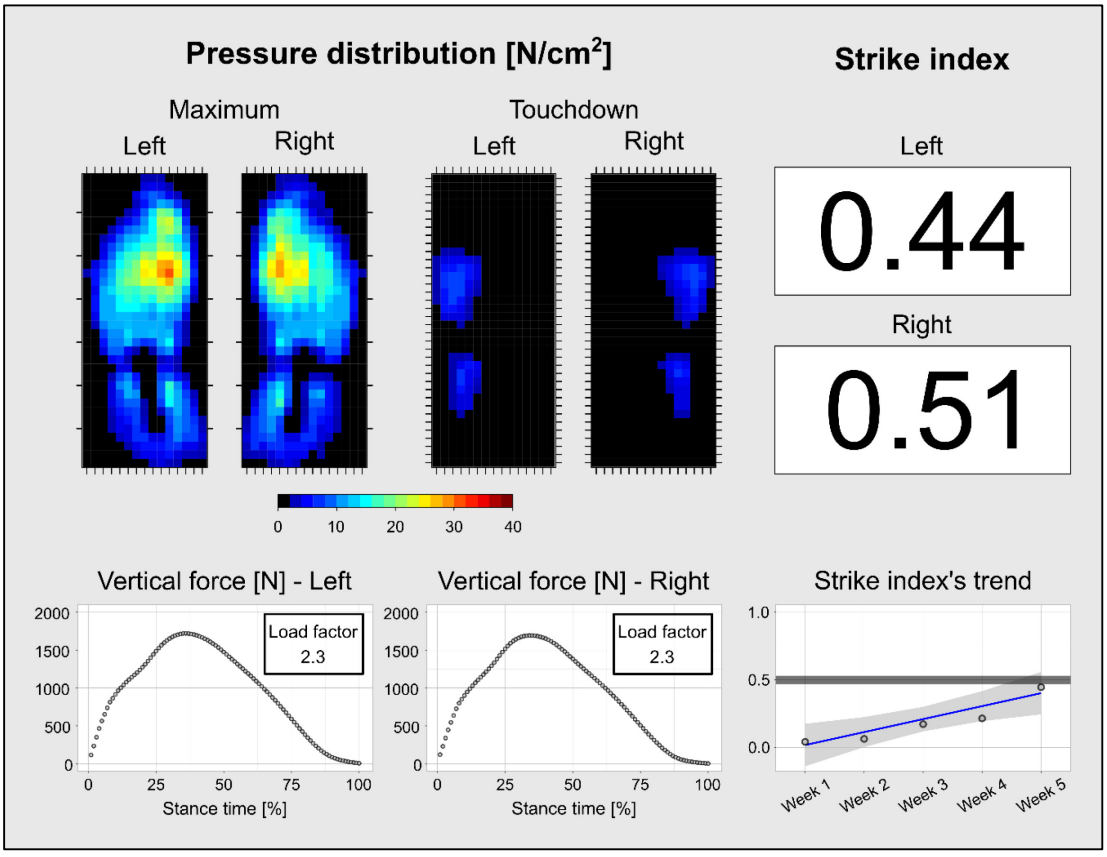

Figure 29 Exemplar output obtained from the custom feedback algorithm that automatically assesses foot strike patterns using plantar pressure distribution data. In this typical screenshot, the plantar pressure distribution (maximum and at touchdown), strike index and vertical ground reaction force values are displayed. The software, written in R version 3.4.3 (R Foundation for Statistical Computing, R Core Team, Vienna, Austria), is being currently used for providing feedback-based foot strike pattern training. In the right-bottom corner, the participants can follow their progress towards the targeted strike index (which is, in this case, 0.5).

The methods developed in the first study presented in this thesis are currently being applied in a number of projects. We are currently investigating the effects of a training intervention, focused on changing the FSP, on the running economy, the modular organisation of motion and on the muscle-tendon unit interaction and energetics. As shown in Figure 29, the first study led to the creation of a powerful feedback tool that has already been used, amongst others, in an intervention study. We trained a group of amateur runners to switch from a rearfoot to a midfoot strike pattern. The participants would come to the laboratory and receive a personalised report on the evolution of their FSP once a week and for 14 weeks. Currently, we are evaluating the effects of this 14-week intervention on the neuromuscular output, muscle-tendon unit properties, local dynamic stability and energy cost of running. Now the software is being used in other two studies based on the analysis of FSPs[313] and in all those projects that require an automatic assessment of gait parameters based on plantar pressure distribution data.

6.2 Muscle synergies as a multidisciplinary tool

The strive for gaining insights into the neuronal mechanisms that control locomotion is not an effort for its own sake. First of all, it is an extremely complicated multidisciplinary exercise that often leads to the integration of theories coming from very different fields, offering the opportunity to unify them[66]. Moreover, a deeper understanding of how living biological objects function is necessary to create or modify existing tools that help us increase our quality of life (e.g. to tackle the mechanisms of those human diseases that affect locomotion or design devices aimed to aid human movement). The concept of muscle synergies has the wonderful potential of being extremely multidisciplinary. By measuring the periphery (i.e. muscle activations) we aim to describe the centre (i.e. the CNS) and the simplicity of the method has been proven to wonderfully adapt to a variety of different fields. During the completion of this doctoral work, I had the chance to interact with scientists having very different backgrounds. In the next subparagraphs I will elaborate on the two most important applications which describe well my future line of research. In the next years I will be using muscle synergies as an important additional tool in projects centred around the neurophysiology of mammalian movement and the design of powered exoskeletons for humans.

6.2.1 Muscle synergies for neurophysiology

It is fairly straightforward to acknowledge the importance of studying locomotion not only on even, solid surfaces, but on uneven terrains as well[126,127,153]. Daily life locomotion includes a variegated mixture of external perturbations that can, in a way or another, challenge the production and control of movements. Muscle synergies would be a clever way for the CNS to simplify movement, but it is highly unlikely that motor outputs could be produced with just a handful of fixed motor modules and primitives [58]. The goal to maintain functionality (e.g. moving effectively from the office to the bus station without falling) is satisfied by the CNS also thanks to the continuous integration of sensorimotor information[190]. This uninterrupted elaboration of inputs and creation of coherent outputs, must be achieved quickly enough also in order to maintain dynamic stability[168,179]. Some of the fundamental links in the integration circuit are the two main functional classes of proprioceptors: muscle spindles GTOs[190]. In humans, proprioception can be permanently impaired due to genetic conditions, viral infections or presbypropria (physiological ageing of proprioception)[314,315]. Akay and his colleagues could show, in the recent past, that the elimination of proprioceptive feedback from muscle spindles in Egr3 (early growth response 3) knockout mice and of muscle spindles and GTOs in mutant mice, produces strong degradation of walking and swimming patterns[190]. However, the extraction of muscle synergies from mice or rat EMG data is an extremely rarely used tool[131]. A collaboration with Prof. Akay started around one year ago and we already managed to extract muscle synergies from the Egr3 knockout mice EMG data. Other groups acknowledged the challenges, highlighted by the second study presented in this thesis, of using NMF to extract synergies from human data and used, amongst others, this publication to either motivate their choices or to describe the potential sources of errors when applying the method[128,316,317]. The short-term goal is to publish a comparison between the EMG activities during walking and swimming of the Egr3 knockout and wild type mice. The goal will be twofold: 1) demonstrating the feasibility of muscle synergies extraction from mouse data and 2) juxtapose an important new approach to the consolidated methods (electrophysiology, motion analysis and mouse genetics) already in use in the lab of Prof. Akay. We hypothesize that the use of muscle synergies is going to be a fundamental tool for improving our knowledge on how the sensory feedback is employed by the CNS in the generation and tuning of locomotor patterns[2,190].

6.2.2 Muscle synergies for exoskeletons

Unsteady locomotion is a challenge that particularly constraints the daily-life of the elderly and pathological population, especially due to fall risk and possible related injuries. Moreover, physical and cognitive degeneration often result in a deterioration of quality of life due to loss of independence. A relatively new branch of robotics has brought, in the last decade, to the perfection of active exoskeletons thought to aid human movement[318]. The torques needed to generate a certain movement can be computed, but the complexity of motion's equations increases dramatically with the number of degrees of freedom[57]. Thus, synergies might be a clever way to store approximate yet sufficient information to build motor commands[57]. Also, given that independent locomotion is the main focus, the challenge of controlling these devices grows exponentially when dealing with uneven grounds[319]. The possibility to implement the muscle synergies concept in the control of a powered exoskeleton has been explored by a few groups in the recent past[319–322]. The potential aim of a synergy-based exoskeleton would be to simplify the control of the device to such an extent that dealing with random external perturbation would become faster. The concept of muscle synergies is a relatively new idea and it has been applied for less than 20 years[57]. However, a broad integration with the world of exoskeletons has to date not yet been achieved. Stefano Rossi and colleagues conceived a compliant lower limb multi-joint exoskeleton for the rehabilitation of ankle knee mobility and locomotion of paediatric patients with neurological diseases[323,324]. The major advantage of using exoskeletons instead of other neuro-robotic techniques (e.g. robotic-assisted gait training) in rehabilitation and habilitation is the great flexibility[318]. Patients wearing an exoskeleton can walk and move autonomously for long periods of time and can thus perform daily-life activities, possibly on a wide range of surfaces. However, there are no studies evaluating the effectiveness of exoskeletons on the field, and hence there is no evidence of their effectiveness in everyday living environments (e.g. for walking on uneven surfaces). We started a collaboration with Prof. Rossi's group aiming to: a) identifying, through muscle synergies, typical movement patterns over uneven surfaces (fundamental set of information for the design of a control system able to adaptively adjust joint impedance and selectively provide force/torque assistance); b) simplify the control of the exoskeleton for adapting the mechatronic system to daily-life requirements of human locomotion.

References

1. Flourens, M. J. P. *Recherches expérimentales sur les propriétés et les fonctions du système nerveux, dans les animaux vertébrés*. (Crevot, 1824).

2. Guertin, P. A. Central Pattern Generator for Locomotion: Anatomical, Physiological, and Pathophysiological Considerations. *Front. Neurol.* **3,** 1–15 (2013).

3. Marder, E. & Calabrese, R. L. Principles of rhythmic motor pattern generation. *Physiol. Rev.* **76,** 687–717 (1996).

4. Selverston, A. I. Invertebrate central pattern generator circuits. *Philos. Trans. R. Soc. B Biol. Sci.* **365,** 2329–2345 (2010).

5. Kiehn, O. & Butt, S. J. B. Physiological, anatomical and genetic identification of CPG neurons in the developing mammalian spinal cord. *Prog. Neurobiol.* **70,** 347–361 (2003).

6. Stuart, D. G. & Hultborn, H. Thomas Graham Brown (1882–1965), Anders Lundberg (1920–), and the neural control of stepping. *Brain Res. Rev.* **59,** 74–95 (2008).

7. Goulding, M. Circuits controlling vertebrate locomotion: moving in a new direction. *Nat. Rev. Neurosci.* **10,** 507–518 (2009).

8. McLean, D. L. & Dougherty, K. J. Peeling back the layers of locomotor control in the spinal cord. *Curr. Opin. Neurobiol.* **33,** 63–70 (2015).

9. Bernstein, N. A. *The co-ordination and regulation of movements*. *Pergamon Press* (Pergamon Press Ltd., 1967).

10. Jankowska, E., Jukes, M. G. M., Lund, S. & Lundberg, A. The Effect of DOPA on the Spinal Cord 5. Reciprocal organization of pathways transmitting excitatory action to alpha motoneurones of flexors and extensors. *Acta Physiol. Scand.* **70,** 369–388 (1967).

11. Jankowska, E., Jukes, M. G. M., Lund, S. & Lundberg, A. The Effect of DOPA on the Spinal Cord 6. Half-centre organization of interneurones transmitting effects from the flexor reflex afferents. *Acta Physiol. Scand.* **70,** 389–402 (1967).

12. Sherrington, C. S. *The integrative action of the nervous system*. (Yale University Press, 1906).

13. Freusberg, A. Reflexbewegungen beim Hunde. *Pflüger, Arch. für die Gesammte Physiol. des Menschen und der Thiere* **9,** 358–391 (1874).

14. Sherrington, C. S. Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing. *J. Physiol.* **40,** 28–121 (1910).

15. Graham Brown, T. On the nature of the fundamental activity of the nervous centres; together with an analysis of the conditioning of rhythmic activity in progression, and a theory of the evolution of function in the nervous system. *J. Physiol.* **48,** 18–46 (1914).

16. Wilson, D. M. & Wyman, R. J. Motor Output Patterns during Random and Rhythmic Stimulation of Locust Thoracic Ganglia. *Biophys. J.* **5,** 121–143 (1965).

17. Wilson, D. M. The central nervous control of flight in a locust. *J. Exp. Biol.* **38,** 471–490 (1961).

18. Holmes Bullock, T. The Origins of Patterned Nervous Discharge. *Behaviour* **17,** 48–58 (1961).

19. Marder, E. & Bucher, D. Central pattern generators and the control of rythmic movements. *Curr. Biol.* **11,** R986–R996 (2001).

20. Robertson, R. M. & Pearson, K. G. Neural circuits in the flight system of the locust. *J. Neurophysiol.* **53,** 110–28 (1985).

21. Ryckebusch, S. & Laurent, G. Rhythmic patterns evoked in locust leg motor neurons by the muscarinic agonist pilocarpine. *J. Neurophysiol.* **69,** 1583–95 (1993).

22. Büschges, A., Akay, T., Gabriel, J. P. & Schmidt, J. Organizing network action for locomotion: Insights from studying insect walking. *Brain Res. Rev.* **57,** 162–171 (2008).

23. Buschges, A., Ludwar, B. C., Bucher, D., Schmidt, J. & DiCaprio, R. A. Synaptic drive contributing to rhythmic activation of motoneurons in the deafferented stick insect walking system. *Eur. J. Neurosci.* **19,** 1856–1862 (2004).

24. Getting, P. A. Neuronal organization of escape swimming in Tritonia. *J. Comp. Physiol. A* **121,** 325–342 (1977).

25. Arshavsky, Y. I., Orlovsky, G. N., Panchin, Y. V., Roberts, A. & Soffe, S. R. Neuronal control of swimming locomotion: analysis of the pteropod mollusc Clione and embryos of the amphibian Xenopus. *Trends Neurosci.* **16,** 227–233 (1993).

26. Hooper, S. L. & DiCaprio, R. A. Crustacean Motor Pattern Generator Networks. *Neurosignals* **13,** 50–69 (2004).

27. Hughes, G. M. & Wiersma, C. A. G. The Co-ordination of Swimmeret Movements in the Crayfish, Procambarus Clarkii (Girard). *J. Exp. Biol.* **37,** 657 (1960).

28. Stein, P. S. G. Intersegmental coordination of swimmeret motoneuron activity in crayfish. *J. Neurophysiol.* **34,** 310–8 (1971).

29. Chrachri, A. & Clarac, F. Induction of rhythmic activity in motoneurons of crayfish thoracic ganglia by cholinergic agonists. *Neurosci. Lett.* **77,** 49–54 (1987).

30. Johnston, R. & Levine, R. Thoracic leg motoneurons in the isolated CNS of adult Manduca produce patterned activity in response to pilocarpine, which is distinct from that produced in larvae. *Invertebr. Neurosci.* **4,** 175–192 (2002).

31. Hückesfeld, S., Schoofs, A., Schlegel, P., Miroschnikow, A. & Pankratz, M. J. Localization of Motor Neurons and Central Pattern Generators for Motor Patterns Underlying Feeding Behavior in Drosophila Larvae. *PLoS One* **10,** e0135011 (2015).

32. Cattaert, D. & Birman, S. Blockade of the central generator of locomotor rhythm by noncompetitive NMDA receptor antagonists in Drosophila larvae. *J. Neurobiol.* **48,** 58–73 (2001).

33. Kohsaka, H., Guertin, P. A. & Nose, A. Neural Circuits Underlying Fly Larval Locomotion. *Curr. Pharm. Des.* **23,** 1722–1733 (2017).

34. Pearson, K. G. & Iles, J. F. Discharge patterns of coxal levator and depressor motoneurones of the cockroach, Periplaneta americana. *J. Exp. Biol.* **52,** 139–65 (1970).

35. Zill, S. N. A model of pattern generation of cockroach walking reconsidered. *J. Neurobiol.* **17,** 317–328 (1986).

36. Guertin, P. A. The mammalian central pattern generator for locomotion. *Brain Res. Rev.* **62,** 45–56 (2009).

37. Grillner, S. & Zangger, P. On the central generation of locomotion in the low spinal cat. *Exp. brain Res.* **34,** 241–61 (1979).

38. Grillner, S. & Wallen, P. Central Pattern Generators for Locomotion, with Special Reference to Vertebrates. *Annu. Rev. Neurosci.* **8,** 233–261 (1985).

39. Forssberg, H., Grillner, S. & Rossignol, S. Phase dependent reflex reversal during walking in chronic spinal cats. *Brain Res.* **85,** 103–107 (1975).

40. Pearson, K. G. & Rossignol, S. Fictive motor patterns in chronic spinal cats. *J. Neurophysiol.* **66,** 1874–87 (1991).

41. Nishimaru, H. & Kudo, N. Formation of the central pattern generator for locomotion in the rat and mouse. *Brain Res. Bull.* **53,** 661–669 (2000).

42. Levine, A. J. *et al.* Identification of a cellular node for motor control pathways. *Nat. Neurosci.* **17,** 586–593 (2014).

43. Roberts, A., Hill, N. a & Hicks, R. Simple mechanisms organise orientation of escape swimming in embryos and hatchling tadpoles of Xenopus laevis. *J. Exp. Biol.* **203,** 1869–1885 (2000).

44. Giszter, S. F., Mussa-Ivaldi, F. A. & Bizzi, E. Convergent force fields organized in the frog's spinal cord. *J. Neurosci.* **13,** 467–491 (1993).

45. Saltiel, P., Tresch, M. C. & Bizzi, E. Spinal cord modular organization and rhythm generation: an NMDA iontophoretic study in the frog. *J. Neurophysiol.* **80,** 2323–2339 (1998).

46. Stein, P. S. G. & Daniels-McQueen, S. Modular organization of turtle spinal interneurons during normal and deletion fictive rostral scratching. *J. Neurosci.* **22,** 6800–9 (2002).

47. Stein, P. S. G. & Grossman, M. L. Central program for scratch reflex in turtle. *J. Comp. Physiol. A* **140,** 287–294 (1980).

48. Grillner, S. On the generation of locomotion in the spiny dogfish. *Exp. Brain Res.* **20,** 459–470 (1974).

49. Grillner, S. *et al.* Neural networks that co-ordinate locomotion and body orientation in lamprey. *Trends Neurosci.* **18,** 270–279 (1995).

50. Grillner, S. & Wallén, P. How does the Lamprey Central Nervous System make the Lamprey Swim? *J. Exp. Biol.* **112,** 337–357 (1984).

51. Bizzi, E., Mussa-Ivaldi, F. A. & Giszter, S. F. Computations underlying the execution of movement: a biological perspective. *Science (80-.).* **253,** 287–291 (1991).

52. Bizzi, E., Cheung, V. C.-K., D'Avella, A., Saltiel, P. & Tresch, M. C. Combining modules for movement. *Brain Res. Rev.* **57,** 125–33 (2008).

53. Ivanenko, Y. P., Poppele, R. E. & Lacquaniti, F. Motor Control Programs and Walking. *Neuroscientist* **12,** 339–348 (2006).

54. Lee, W. A. Neuromotor Synergies as a Basis for Coordinated Intentional Action. *J. Mot. Behav.* **16,** 135–170 (1984).

55. Tresch, M. C., Saltiel, P. & Bizzi, E. The construction of movement by the spinal cord. *Nat. Neurosci.* **2,** 162–167 (1999).

56. Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401,** 788–91 (1999).

57. D'Avella, A. in *Progress in Motor Control* (ed. Sternad, D.) **629,** 3–19 (Springer US, 2016).

58. Bizzi, E. & Cheung, V. C.-K. The neural origin of muscle synergies. *Front. Comput. Neurosci.* **7,** 51 (2013).

59. Ting, L. H. *et al.* Neuromechanical Principles Underlying Movement Modularity and Their Implications for Rehabilitation. *Neuron* **86,** 38–54 (2015).

60. Bizzi, E. & Ajemian, R. J. A Hard Scientific Quest: Understanding Voluntary Movements. *Daedalus* **144,** 83–95 (2015).

61. Latash, M. L., Scholz, J. P. & Schöner, G. Toward a new theory of motor synergies. *Motor Control* **11,** 276–308 (2007).

62. Ting, L. H. & McKay, J. L. Neuromechanics of muscle synergies for posture and movement. *Curr. Opin. Neurobiol.* **17,** 622–8 (2007).

63. Tresch, M. C. & Jarc, A. The case for and against muscle synergies. *Curr. Opin. Neurobiol.* **19,** 601–7 (2009).

64. d'Avella, A., Giese, M., Ivanenko, Y. P., Schack, T. & Flash, T. Editorial: Modularity in motor control: from muscle synergies to cognitive action representation. *Front. Comput. Neurosci.* **9,** 1–6 (2015).

65. Damiano, D. Muscle synergies: input or output variables for neural control? *Dev. Med.*

Child Neurol. **57,** 1091–1092 (2015).

66. Latash, M. L., Levin, M. F., Scholz, J. P. & Schöner, G. Motor Control Theories and Their Applications. *Med.* **46,** 382–392 (2010).

67. Turvey, M. T. & Fonseca, S. Nature of Motor Control: Perspectives and Issues. *Adv. Exp. Med. Biol.* **629,** 93–123 (2009).

68. Flash, T. & Hochner, B. Motor primitives in vertebrates and invertebrates. *Curr. Opin. Neurobiol.* **15,** 660–666 (2005).

69. Lacquaniti, F., Ivanenko, Y. P. & Zago, M. Patterned control of human locomotion. *J. Physiol.* **590,** 2189–2199 (2012).

70. Giszter, S. F. Motor primitives-new data and future questions. *Curr. Opin. Neurobiol.* **33,** 156–165 (2015).

71. Cheung, V. C.-K. Sensory modulation of muscle synergies for motor adaptation during natural behaviors. (Harvard University, 2007).

72. Giszter, S. F. & Hart, C. B. Motor primitives and synergies in the spinal cord and after injury-the current state of play. *Ann. N. Y. Acad. Sci.* **1279,** 114–126 (2013).

73. Harris, C. M. & Wolpert, D. M. Signal-dependent noise determines motor planning. *Nature* **394,** 780–784 (1998).

74. Tresch, M. C., Cheung, V. C.-K. & D'Avella, A. Matrix factorization algorithms for the identification of muscle synergies: evaluation on simulated and experimental data sets. *J. Neurophysiol.* **95,** 2199–212 (2006).

75. Santuz, A., Ekizos, A., Janshen, L., Baltzopoulos, V. & Arampatzis, A. On the Methodological Implications of Extracting Muscle Synergies from Human Locomotion. *Int. J. Neural Syst.* **27,** 1750007 (2017).

76. Lee, D. D. & Seung, H. S. Algorithms for Non-negative Matrix Factorization. in

Advances in Neural Information Processing Systems 556–562 (The MIT Press, 2001).

77. Santuz, A., Ekizos, A., Janshen, L., Baltzopoulos, V. & Arampatzis, A. The Influence of Footwear on the Modular Organization of Running. *Front. Physiol.* **8,** 958 (2017).

78. Devarajan, K. & Cheung, V. C.-K. On nonnegative matrix factorization algorithms for signal-dependent noise with application to electromyography data. *Neural Comput.* **26,** 1128–68 (2014).

79. Cheung, V. C.-K., D'Avella, A., Tresch, M. C. & Bizzi, E. Central and sensory contributions to the activation and organization of muscle synergies during natural motor behaviors. *J. Neurosci.* **25,** 6419–34 (2005).

80. D'Avella, A. & Bizzi, E. Shared and specific muscle synergies in natural motor behaviors. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 3076–81 (2005).

81. Gao, Z., Zhai, G. & Wang, J. Spatially-weighted nonnegative matrix factorization with application to temporal psychovisual modulation. *Digit. Signal Process.* **67,** 123–130 (2017).

82. Wang, J. J.-Y. & Gao, X. Max–min distance nonnegative matrix factorization. *Neural Networks* **61,** 75–84 (2015).

83. Hu, W., Choi, K.-S., Wang, P., Jiang, Y. & Wang, S. Convex nonnegative matrix factorization with manifold regularization. *Neural Networks* **63C,** 94–103 (2014).

84. Devarajan, K., Wang, G. & Ebrahimi, N. A unified statistical approach to non-negative matrix factorization and probabilistic latent semantic indexing. *Mach. Learn.* **99,** 137–163 (2015).

85. Kim, J., He, Y. & Park, H. Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework. *J. Glob. Optim.* **58,** 285–319 (2013).

86. Cichocki, A., Zdunek, R. & Amari, S. Csiszár's Divergences for Non-negative Matrix

Factorization: Family of New Algorithms. in *Independent Component Analysis and Blind Signal Separation, 6th International Conference, ICA 2006* (eds. Rosca, J., Erdogmus, D., Príncipe, J. C. & Haykin, S.) **3889,** 32–39 (Springer Berlin Heidelberg, 2006).

87. Blondel, V. D., Ngoc-Diep, H. & van Dooren, P. Weighted nonnegative matrix factorization and face feature extraction. *Submitt. to Image Vis. ...* (2007).

88. Li, X. & Fukui, K. Fisher Non-negative Matrix Factorization with Pairwise Weighting. *Proc. IAPR Conf. Mach. Vis. Appl.* 380–383 (2007).

89. Févotte, C., Bertin, N. & Durrieu, J.-L. Nonnegative matrix factorization with the Itakura-Saito divergence: with application to music analysis. *Neural Comput.* **21,** 793–830 (2009).

90. Dominici, N. *et al.* Locomotor primitives in newborn babies and their development. *Science (80-.).* **334,** 997–9 (2011).

91. Cappellini, G., Ivanenko, Y. P., Poppele, R. E. & Lacquaniti, F. Motor patterns in human walking and running. *J. Neurophysiol.* **95,** 3426–37 (2006).

92. Oliveira, A. S. C., Gizzi, L., Farina, D. & Kersting, U. G. Motor modules of human locomotion: influence of EMG averaging, concatenation, and number of step cycles. *Front. Hum. Neurosci.* **8,** 335 (2014).

93. Saito, A., Tomita, A., Ando, R., Watanabe, K. & Akima, H. Similarity of muscle synergies extracted from the lower limb including the deep muscles between level and uphill treadmill walking. *Gait Posture* **59,** 134–139 (2018).

94. Shuman, B. R., Schwartz, M. H. & Steele, K. M. Electromyography Data Processing Impacts Muscle Synergies during Gait for Unimpaired Children and Children with Cerebral Palsy. *Front. Comput. Neurosci.* **11,** 1–9 (2017).

95. Janshen, L., Santuz, A., Ekizos, A. & Arampatzis, A. Modular control during incline and level walking in humans. *J. Exp. Biol.* **220,** 807–813 (2017).

96. Buurke, T. J. W., Lamoth, C. J. C., van der Woude, L. H. V. & den Otter, A. R. Synergistic structure in the speed dependent modulation of muscle activity in human walking. *PLoS One* **11,** 1–19 (2016).

97. Gui, K. & Zhang, D. Influence of locomotion speed on biomechanical subtask and muscle synergy. *J. Electromyogr. Kinesiol.* **30,** 209–215 (2016).

98. Pérez-Nombela, S. *et al.* Modular control of gait after incomplete spinal cord injury: differences between sides. *Spinal Cord* 1–8 (2016). doi:10.1038/sc.2016.99

99. Lencioni, T. *et al.* Are Modular Activations Altered in Lower Limb Muscles of Persons with Multiple Sclerosis during Walking? Evidence from Muscle Synergies and Biomechanical Analysis. *Front. Hum. Neurosci.* **10,** 1–14 (2016).

100. Yokoyama, H., Ogawa, T., Kawashima, N., Shinya, M. & Nakazawa, K. Distinct sets of locomotor modules control the speed and modes of human locomotion. *Sci. Rep.* **6,** 1–14 (2016).

101. Meyer, A. J. *et al.* Muscle Synergies Facilitate Computational Prediction of Subject-Specific Walking Motions. *Front. Bioeng. Biotechnol.* **4,** (2016).

102. Kim, Y., Bulea, T. C. & Damiano, D. L. Novel Methods to Enhance Precision and Reliability in Muscle Synergy Identification during Walking. *Front. Hum. Neurosci.* **10,** 455 (2016).

103. Coscia, M. *et al.* Muscle synergies and spinal maps are sensitive to the asymmetry induced by a unilateral stroke. *J. Neuroeng. Rehabil.* **12,** 39 (2015).

104. Hagio, S., Fukuda, M. & Kouzaki, M. Identification of muscle synergies associated with gait transition in humans. *Front. Hum. Neurosci.* **9,** 48 (2015).

105. Gonzalez-Vargas, J. *et al.* A predictive model of muscle excitations based on muscle modularity for a large repertoire of human locomotion conditions. *Front. Comput. Neurosci.* **9,** 1–14 (2015).

106. Martino, G. *et al.* Neuromuscular adjustments of gait associated with unstable conditions. *J. Neurophysiol.* **114,** jn.00029.2015 (2015).

107. Tang, L. *et al.* Muscle synergy analysis in children with cerebral palsy. *J. Neural Eng.* **12,** 046017 (2015).

108. Nazifi, M. M., Yoon, H. U., Beschorner, K. & Hur, P. Shared and task-specific muscle synergies during normal walking and slipping. *Am. Soc. Biomech.* **11,** 1–14 (2015).

109. Licence, S., Smith, R., McGuigan, M. P. & Earnest, C. P. Gait pattern alterations during walking, texting and walking and texting during cognitively distractive tasks while negotiating common pedestrian obstacles. *PLoS One* **10,** 1–11 (2015).

110. Maclellan, M. J. *et al.* Muscle activation patterns are bilaterally linked during split-belt treadmill walking in humans. *J. Neurophysiol.* **111,** 1541–52 (2014).

111. Barroso, F. O. *et al.* Shared muscle synergies in human walking and cycling. *J. Neurophysiol.* jn.00220.2014- (2014). doi:10.1152/jn.00220.2014

112. Routson, R. L., Kautz, S. A. & Neptune, R. R. Modular organization across changing task demands in healthy and poststroke gait. *Physiol. Rep.* **2,** 1–14 (2014).

113. Chvatal, S. A. & Ting, L. H. Common muscle synergies for balance and walking. *Front. Comput. Neurosci.* **7,** 14 (2013).

114. Rodriguez, K. L., Roemmich, R. T., Cam, B., Fregly, B. J. & Hass, C. J. Persons with Parkinson's disease exhibit decreased neuromuscular complexity during gait. *Clin. Neurophysiol.* **124,** 1390–1397 (2013).

115. Allen, J. L. & Neptune, R. R. Three-dimensional modular control of human walking. *J. Biomech.* **45,** 2157–63 (2012).

116. Oliveira, A. S. C., Gizzi, L., Kersting, U. G. & Farina, D. Modular organization of balance control following perturbations during walking. *J. Neurophysiol.* **108,** 1895–906 (2012).

117. Chvatal, S. A. & Ting, L. H. Voluntary and reactive recruitment of locomotor muscle synergies during perturbed walking. *J. Neurosci.* **32,** 12237–50 (2012).

118. Bolton, D. A. E. & Misiaszek, J. E. Compensatory balance reactions during forward and backward walking on a treadmill. *Gait Posture* **35,** 681–4 (2012).

119. Clark, D. J., Ting, L. H., Zajac, F. E., Neptune, R. R. & Kautz, S. A. Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke. *J. Neurophysiol.* **103,** 844–57 (2010).

120. McGowan, C. P., Neptune, R. R., Clark, D. J. & Kautz, S. A. Modular control of human walking: Adaptations to altered mechanical demands. *J. Biomech.* **43,** 412–419 (2010).

121. Courtine, G., Papaxanthis, C. & Schieppati, M. Coordinated modulation of locomotor muscle synergies constructs straight-ahead and curvilinear walking in humans. *Exp. Brain Res.* **170,** 320–35 (2006).

122. Ivanenko, Y. P., Poppele, R. E. & Lacquaniti, F. Five basic muscle activation patterns account for muscle activity during human locomotion. *J. Physiol.* **556,** 267–282 (2004).

123. Allen, J. L., McKay, J. L., Sawers, A., Hackney, M. E. & Ting, L. H. Increased neuromuscular consistency in gait and balance after partnered, dance-based rehabilitation in Parkinson's disease. *J. Neurophysiol.* **118,** 363–373 (2017).

124. Nishida, K., Hagio, S., Kibushi, B., Moritani, T. & Kouzaki, M. Comparison of muscle synergies for running between different foot strike patterns. *PLoS One* **12,** e0171535 (2017).

125. Burfoot, A. The history of the marathon: 1976-Present. *Sport. Med.* **37,** 284–287 (2007).

126. Santuz, A., Ekizos, A., Eckardt, N., Kibele, A. & Arampatzis, A. Challenging human locomotion: stability and modular organisation in unsteady conditions. *Sci. Rep.* **8,**

2740 (2018).

127. Biewener, A. A. & Daley, M. A. Unsteady locomotion: integrating muscle function with whole body dynamics and neuromuscular control. *J. Exp. Biol.* **210,** 2949–60 (2007).

128. Banks, C. L., Pai, M. M., McGuirk, T. E., Fregly, B. J. & Patten, C. Methodological Choices in Muscle Synergy Analysis Impact Differentiation of Physiological Characteristics Following Stroke. *Front. Comput. Neurosci.* **11,** 1–12 (2017).

129. Shuman, B. *et al.* Repeatability of muscle synergies within and between days for typically developing children and children with cerebral palsy. *Gait Posture* **45,** 127–132 (2016).

130. Falaki, A., Huang, X., Lewis, M. M. & Latash, M. L. Impaired synergic control of posture in Parkinson's patients without postural instability. *Gait Posture* **44,** 209–215 (2016).

131. Wenger, N. *et al.* Spatiotemporal neuromodulation therapies engaging muscle synergies improve motor control after spinal cord injury. *Nat. Med.* **22,** 5–7 (2016).

132. Latash, M. L. & Anson, J. G. Synergies in health and disease: Relations to adaptive changes in motor coordination. *Phys. Ther.* **86,** 1151–60 (2006).

133. Rana, M., Yani, M. S., Asavasopon, S., Fisher, B. E. & Kutch, J. J. Brain Connectivity Associated with Muscle Synergies in Humans. *J. Neurosci.* **35,** 14708–14716 (2015).

134. Amundsen Huffmaster, S. L., Van Acker III, G. M., Luchies, C. W. & Cheney, P. D. Muscle synergies obtained from comprehensive mapping of the primary motor cortex forelimb representation using high-frequency long-duration ICMS. *J. Neurophysiol.* jn.00784.2016 (2017). doi:10.1152/jn.00784.2016

135. Overduin, S. A., D'Avella, A., Roh, J., Carmena, J. M. & Bizzi, E. Representation of Muscle Synergies in the Primate Brain. *J. Neurosci.* **35,** 12615–12624 (2015).

136. Zampieri, N., Jessell, T. M. & Murray, A. J. Mapping Sensory Circuits by Anterograde Transsynaptic Transfer of Recombinant Rabies Virus. *Neuron* **81,** 766–778 (2014).

137. Wickersham, I. R., Finke, S., Conzelmann, K. K. & Callaway, E. M. Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* **4,** 47–49 (2007).

138. van Ingen Schenau, G. J. Some fundamental aspects of the biomechanics of overground versus treadmill locomotion. *Med. Sci. Sports Exerc.* **12,** 257–61 (1980).

139. Jones, A. M. & Doust, J. H. A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. *J. Sports Sci.* **14,** 321–327 (1996).

140. Mooses, M., Tippi, B., Mooses, K., Durussel, J. & Mäestu, J. Better economy in field running than on the treadmill: evidence from high-level distance runners. *Biol. Sport* **32,** 155–159 (2014).

141. Parvataneni, K., Ploeg, L., Olney, S. J. & Brouwer, B. Kinematic, kinetic and metabolic parameters of treadmill versus overground walking in healthy older adults. *Clin. Biomech.* **24,** 95–100 (2009).

142. Warabi, T., Kato, M., Kiriyama, K., Yoshida, T. & Kobayashi, N. Treadmill walking and overground walking of human subjects compared by recording sole-floor reaction force. *Neurosci. Res.* **53,** 343–8 (2005).

143. Lee, S. J. & Hidler, J. Biomechanics of overground vs. treadmill walking in healthy individuals. *J. Appl. Physiol.* **104,** 747–55 (2008).

144. Chia, L. C., Licari, M. K., Guelfi, K. J. & Reid, S. L. Investigation of treadmill and overground running: implications for the measurement of oxygen cost in children with developmental coordination disorder. *Gait Posture* **40,** 464–70 (2014).

145. Riley, P. O. *et al.* A kinematics and kinetic comparison of overground and treadmill running. *Med. Sci. Sport. Exerc.* **40,** 1093–100 (2008).

146. Bailey, J., Mata, T. & Mercer, J. A. Is the Relationship Between Stride Length,

Frequency, and Velocity Influenced by Running on a Treadmill or Overground? *Int. J. Exerc. Sci.* **10,** 1067–1075 (2017).

147. Oliveira, A. S. C., Gizzi, L., Ketabi, S., Farina, D. & Kersting, U. G. Modular Control of Treadmill vs Overground Running. *PLoS One* **11,** e0153307 (2016).

148. Iosa, M., Gizzi, L., Tamburella, F. & Dominici, N. Editorial: Neuro-motor control and feed-forward models of locomotion in humans. *Front. Hum. Neurosci.* **9,** 1–4 (2015).

149. Jordan, K., Challis, J. H., Cusumano, J. P. & Newell, K. M. Stability and the time-dependent structure of gait variability in walking and running. *Hum. Mov. Sci.* **28,** 113–128 (2009).

150. Cavanagh, P. R., Pollock, M. L. & Landa, J. A Biomechanical Comparison of Elite and Good Distance Runners. *Ann. New York Acad. Sci.* **301,** 328–345 (1977).

151. Carrier, D. R., Heglund, N. C. & Earls, K. D. Variable gearing during locomotion in the human musculoskeletal system. *Science (80-.).* **265,** 651–653 (1994).

152. Lieberman, D. E. Human locomotion and heat loss: An evolutionary perspective. *Compr. Physiol.* **5,** 99–117 (2015).

153. Daley, M. A. in *Understanding Mammalian Locomotion* 277–306 (John Wiley & Sons, Inc, 2016). doi:10.1002/9781119113713.ch11

154. Cavanagh, P. R. & Williams, K. R. The effect of stride length variation on oxygen uptake during distance running. *Med. Sci. Sport. Exerc.* **14,** 30–35 (1982).

155. Hoyt, D. F. & Taylor, C. R. Gait and the energetics of locomotion in horses. *Nature* **292,** 239–240 (1981).

156. Soule, R. G. & Goldman, R. F. Terrain coefficients for energy cost prediction. *J. Appl. Physiol.* **32,** 706–8 (1972).

157. Zamparo, P., Perini, R., Orizio, C., Sacher, M. & Ferretti, G. The energy cost of

walking or running on sand. *Eur. J. Appl. Physiol. Occup. Physiol.* **65,** 183–187 (1992).

158. Pinnington, H. C. & Dawson, B. The energy cost of running on grass compared to soft dry beach sand. *J. Sci. Med. Sport* **4,** 416–430 (2001).

159. Jensen, K., Johansen, L. & Kärkkäinen, O.-P. Economy in track runners and orienteers during path and terrain running. *J. Sports Sci.* **17,** 945–950 (1999).

160. Heinonen, A. O., Karvonen, M. J. & Ruosteenoja, R. The energy expenditure of walking on snow at various speeds. *Ergonomics* **2,** 389–394 (1959).

161. Glasow, W. & Müller, E. A. Das Gehen auf verschiedenen Böden. *Arbeitsphysiologie* **14,** 319–321 (1951).

162. Pandolf, K. B., Haisman, M. F. & Goldman, R. F. Metabolic Energy Expenditure and Terrain Coefficients for Walking on Snow. *Ergonomics* **19,** 683–690 (1976).

163. Voloshina, A. S., Kuo, A. D., Daley, M. A. & Ferris, D. P. Biomechanics and energetics of walking on uneven terrain. *J. Exp. Biol.* **216,** 3963–70 (2013).

164. Voloshina, A. S. & Ferris, D. P. Biomechanics and energetics of running on uneven terrain. *J. Exp. Biol.* **218,** 711–719 (2015).

165. Ferris, D. P., Louie, M. & Farley, C. T. Running in the real world: adjusting leg stiffness for different surfaces. *Proc. R. Soc. B Biol. Sci.* **265,** 989–994 (1998).

166. Daley, M. A. & Usherwood, J. R. Two explanations for the compliant running paradox: reduced work of bouncing viscera and increased stability in uneven terrain. *Biol. Lett.* **6,** 418–421 (2010).

167. Usherwood, J. R. & Bertram, J. E. A. Understanding brachiation: insight from a collisional perspective. *J. Exp. Biol.* **206,** 1631–1642 (2003).

168. Daley, M. A. & Biewener, A. A. Running over rough terrain reveals limb control for

intrinsic stability. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 15681–6 (2006).

169. Yandell, M. B. & Zelik, K. E. Preferred Barefoot Step Frequency is Influenced by Factors Beyond Minimizing Metabolic Rate. *Sci. Rep.* **6,** 23243 (2016).

170. Hak, L. *et al.* Speeding up or slowing down?: Gait adaptations to preserve gait stability in response to balance perturbations. *Gait Posture* **36,** 260–264 (2012).

171. McAndrew, P. M., Wilken, J. M. & Dingwell, J. B. Dynamic stability of human walking in visually and mechanically destabilizing environments. *J. Biomech.* **44,** 644–649 (2011).

172. Chang, M. D., Sejdić, E., Wright, V. & Chau, T. Measures of dynamic stability: Detecting differences between walking overground and on a compliant surface. *Hum. Mov. Sci.* **29,** 977–986 (2010).

173. Hak, L. *et al.* Stepping strategies for regulating gait adaptability and stability. *J. Biomech.* **46,** 905–911 (2013).

174. Birn-Jeffery, A. V. *et al.* Don't break a leg: running birds from quail to ostrich prioritise leg safety and economy on uneven terrain. *J. Exp. Biol.* **217,** 3786–3796 (2014).

175. Sterzing, T., Apps, C., Ding, R. & Cheung, J. Running on an unpredictable irregular surface changes lower limb biomechanics and subjective perception compared to running on a regular surface. *J. Foot Ankle Res.* **7,** A80 (2014).

176. Ernst, M., Götze, M., Müller, R. & Blickhan, R. Vertical adaptation of the center of mass in human running on uneven ground. *Hum. Mov. Sci.* **38,** 293–304 (2014).

177. Daley, M. A. Biomechanics: Running Over Uneven Terrain Is a No-Brainer. *Curr. Biol.* **18,** R1064–R1066 (2008).

178. Patla, A. E. Strategies for Dynamic Stability During Adaptive Human Locomotion. *Eng. Med. Biol. Mag. IEEE* **22,** 48–52 (2003).

179. Daley, M. A., Usherwood, J. R., Felix, G. & Biewener, A. A. Running over rough terrain: guinea fowl maintain dynamic stability despite a large unexpected change in substrate height. *J. Exp. Biol.* **209,** 171–187 (2006).

180. Kiehn, O. Locomotor Circuits in the Mammalian Spinal Cord. *Annu. Rev. Neurosci.* **29,** 279–306 (2006).

181. Dougherty, K. J. *et al.* Locomotor Rhythm Generation Linked to the Output of Spinal Shox2 Excitatory Interneurons. *Neuron* **80,** 920–933 (2013).

182. Talpalar, A. E. *et al.* Dual-mode operation of neuronal networks involved in left-right alternation. *Nature* **500,** 85–88 (2013).

183. Rybak, I. A., Dougherty, K. J. & Shevtsova, N. A. Organization of the Mammalian Locomotor CPG: Review of Computational Model and Circuit Architectures Based on Genetically Identified Spinal Interneurons. *eNeuro* **2,** (2015).

184. Shevtsova, N. A. & Rybak, I. A. Organization of flexor-extensor interactions in the mammalian spinal cord: insights from computational modelling. *J. Physiol.* **594,** 6117–6131 (2016).

185. Bretzner, F. Contribution of the Motor Cortex to the Structure and the Timing of Hindlimb Locomotion in the Cat: A Microstimulation Study. *J. Neurophysiol.* **94,** 657–672 (2005).

186. Yakovenko, S., Krouchev, N. & Drew, T. Sequential activation of motor cortical neurons contributes to intralimb coordination during reaching in the cat by modulating muscle synergies. *J. Neurophysiol.* **105,** 388–409 (2011).

187. Rathelot, J.-A. & Strick, P. L. Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 918–923 (2009).

188. Griffin, D. M., Hoffman, D. S. & Strick, P. L. Corticomotoneuronal cells are ‘functionally tuned’. *Science (80-.).* **350,** 667–670 (2015).

189. Hägglund, M., Borgius, L., Dougherty, K. J. & Kiehn, O. Activation of groups of excitatory neurons in the mammalian spinal cord or hindbrain evokes locomotion. *Nat. Neurosci.* **13,** 246–252 (2010).

190. Akay, T., Tourtellotte, W. G., Arber, S. & Jessell, T. M. Degradation of mouse locomotor pattern in the absence of proprioceptive sensory feedback. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 16877–82 (2014).

191. Prochazka, A., Gillard, D. & Bennett, D. J. Positive Force Feedback Control of Muscles. *J. Neurophysiol.* **77,** 3226–3236 (1997).

192. Duysens, J., De Groote, F. & Jonkers, I. The flexion synergy, mother of all synergies and father of new models of gait. *Front. Comput. Neurosci.* **7,** 14 (2013).

193. Conway, B. A., Hultborn, H. & Kiehn, O. Proprioceptive input resets central locomotor rhythm in the spinal cat. *Exp. Brain Res.* **68,** 643–656 (1987).

194. Bourane, S. *et al.* Identification of a spinal circuit for light touch and fine motor control. *Cell* **160,** 503–516 (2015).

195. Koch, S. C. *et al.* RORβ Spinal Interneurons Gate Sensory Transmission during Locomotion to Secure a Fluid Walking Gait. *Neuron* **96,** 1419–1431.e5 (2017).

196. Santuz, A., Ekizos, A. & Arampatzis, A. A Pressure Plate-Based Method for the Automatic Assessment of Foot Strike Patterns During Running. *Ann. Biomed. Eng.* **44,** 1646–1655 (2016).

197. Beely, F. Zur Mechanik des Stehens. Ueber die Bedeutung des Fussgewölbes beim Stehen. *Arch. für Klin. Chir.* **27,** 457–471 (1882).

198. Momburg, F. *Der Gang des Menschen und die Fussgeschwulst*. **25,** (Hirschwald, 1908).

199. Elftman, H. A cinematic study of the distribution of pressure in the human foot. *Anat. Rec.* **59,** 481–491 (1934).

200. Cavagna, G. A., Komarek, L. & Mazzoleni, S. The mechanics of sprint running. *J. Neurosci.* **217,** 709–721 (1971).

201. Cavanagh, P. R. & Lafortune, M. A. Ground reaction forces in distance running. *J. Biomech.* **13,** 397–406 (1980).

202. Bramble, D. M. & Lieberman, D. E. Endurance running and the evolution of Homo. *Nature* **432,** 345–52 (2004).

203. Hasegawa, H., Yamauchi, T. & Kraemer, W. J. Foot strike patterns of runners at the 15-km point during an elite-level half marathon. *J. Strength Cond. Res.* **21,** 888–893 (2007).

204. Lieberman, D. E. *et al.* Foot strike patterns and collision forces in habitually barefoot versus shod runners. *Nature* **463,** 531–5 (2010).

205. Daoud, A. I. *et al.* Foot strike and injury rates in endurance runners: a retrospective study. *Med. Sci. Sport. Exerc.* **44,** 1325–34 (2012).

206. Pohl, M. B., Mullineaux, D. R., Milner, C. E., Hamill, J. & Davis, I. S. Biomechanical predictors of retrospective tibial stress fractures in runners. *J. Biomech.* **41,** 1160–5 (2008).

207. Milner, C. E., Ferber, R., Pollard, C. D., Hamill, J. & Davis, I. S. Biomechanical factors associated with tibial stress fracture in female runners. *Med. Sci. Sport. Exerc.* **38,** 323–8 (2006).

208. Zifchock, R. A., Davis, I. S. & Hamill, J. Kinetic asymmetry in female runners with and without retrospective tibial stress fractures. *J. Biomech.* **39,** 2792–7 (2006).

209. Hamill, J. & Gruber, A. H. Running injuries: forefoot versus rearfoot and barefoot versus shod: a biomechanist's perspective. in *30th Annual Conference of Biomechanics in Sports* 64–67 (2012).

210. Larson, P. *et al.* Foot strike patterns of recreational and sub-elite runners in a long-

distance road race. *J. Sports Sci.* **29,** 1665–73 (2011).

211. Di Michele, R. & Merni, F. The concurrent effects of strike pattern and ground-contact time on running economy. *J. Sci. Med. Sport* **17,** 414–418 (2014).

212. Williams, D. S., McClay, I. S. & Manal, K. T. Lower Extremity Mechanics in Runners with a Converted Forefoot Strike Pattern. *J. Appl. Biomech.* **16,** 210–218 (2000).

213. McCallion, C., Donne, B., Fleming, N. & Blanksby, B. Acute Differences in Foot Strike and Spatiotemporal Variables for Shod, Barefoot or Minimalist Male Runners. *J. Sport. Sci. Med.* **13,** 280–286 (2014).

214. Gruber, A. H., Umberger, B. R., Braun, B. & Hamill, J. Economy and rate of carbohydrate oxidation during running with rearfoot and forefoot strike patterns. *J. Appl. Physiol.* **115,** 194–201 (2013).

215. Perl, D. P., Daoud, A. I. & Lieberman, D. E. Effects of footwear and strike type on running economy. *Med. Sci. Sport. Exerc.* **44,** 1335–43 (2012).

216. Logan, S., Hunter, I., Hopkins, J. T., Feland, J. B. & Parcell, A. C. Ground reaction force differences between running shoes, racing flats, and distance spikes in runners. *J. Sport. Sci. Med.* **9,** 147–153 (2010).

217. de Almeida, M. O., Saragiotto, B. T., Yamato, T. P. & Lopes, A. D. Is the rearfoot pattern the most frequently foot strike pattern among recreational shod distance runners? *Phys. Ther. Sport* 1–5 (2014). doi:10.1016/j.ptsp.2014.02.005

218. Kasmer, M. E., Liu, X.-C., Roberts, K. G. & Valadao, J. M. Foot-strike pattern and performance in a marathon. *Int. J. Sports Physiol. Perform.* **8,** 286–92 (2013).

219. Bertelsen, M. L., Jensen, J. F., Nielsen, M. H., Nielsen, R. O. & Rasmussen, S. Footstrike patterns among novice runners wearing a conventional, neutral running shoe. *Gait Posture* 10–12 (2012). doi:10.1016/j.gaitpost.2012.11.022

220. Altman, A. R. & Davis, I. S. A kinematic method for footstrike pattern detection in

barefoot and shod runners. *Gait Posture* **35,** 298–300 (2012).

221. Giandolini, M. *et al.* A simple field method to identify foot strike pattern during running. *J. Biomech.* **47,** 1588–1593 (2014).

222. Boyer, E. R. & Derrick, T. R. Select Injury-Related Variables Are Affected by Stride Length and Foot Strike Style During Running. *Am. J. Sports Med.* **43,** 2310–2317 (2015).

223. Mann, R. *et al.* Reliability and validity of pressure and temporal parameters recorded using a pressure-sensitive insole during running. *Gait Posture* **39,** 455–459 (2014).

224. Pohl, M. B. & Buckley, J. G. Changes in foot and shank coupling due to alterations in foot strike pattern during running. *Clin. Biomech.* **23,** 334–41 (2008).

225. Laughton, C. A., Davis, I. S. & Hamill, J. Effect of Strike Pattern and Orthotic Intervention on Tibial Shock During Running. *J. Appl. Biomech.* **19,** 153–168 (2003).

226. Hatala, K. G., Dingwall, H. L., Wunderlich, R. E. & Richmond, B. G. Variation in Foot Strike Patterns during Running among Habitually Barefoot Populations. *PLoS One* **8,** e52548 (2013).

227. Scholten, S. D. *et al.* Foot strike patterns after obstacle clearance during running. *Med. Sci. Sport. Exerc.* **34,** 123–129 (2002).

228. Lees, A. & Bouracier, J. The longitudinal variability of ground reaction forces in experienced and inexperienced runners. *Ergonomics* **37,** 197–206 (1994).

229. Lun, V., Meeuwisse, W. H., Stergiou, P. & Stefanyshyn, D. Relation between running injury and static lower limb alignment in recreational runners. *Br. J. Sports Med.* **38,** 576–80 (2004).

230. Braunstein, B., Arampatzis, A., Eysel, P. & Brüggemann, G.-P. Footwear affects the gearing at the ankle and knee joints during running. *J. Biomech.* **43,** 2120–5 (2010).

231. Bennell, K. L. *et al.* Risk factors for stress fractures in female track-and-field athletes: a retrospective analysis. *Clin. J. Sport Med.* **5,** 229–35 (1995).

232. Fletcher, J. R., Esau, S. P. & MacIntosh, B. R. Economy of running: beyond the measurement of oxygen uptake. *J. Appl. Physiol.* **107,** 1918–22 (2009).

233. Treutwein, B. Adaptive psychophysical procedures. *Vision Res.* **35,** 2503–2522 (1995).

234. Borg, G. A. V. Psychophysical bases of perceived exertion. *Med. Sci. Sport. Exerc.* **14,** 377–381 (1982).

235. White, S. C., Gilchrist, L. a. & Christina, K. a. Within-day accommodation effects on vertical reaction forces for treadmill running. *J. Appl. Biomech.* **18,** 74–82 (2002).

236. Kernozek, T. W., Meardon, S. & Vannatta, C. N. In-Shoe Loading in Rearfoot and Non-Rearfoot Strikers during Running Using Minimalist Footwear. *Int. J. Sports Med.* (2014). doi:10.1055/s-0034-1372627

237. Chambon, N., Delattre, N., Guéguen, N., Berton, E. & Rao, G. Is midsole thickness a key parameter for the running pattern? *Gait Posture* 2–7 (2014). doi:10.1016/j.gaitpost.2014.02.005

238. Squadrone, R. & Gallozzi, C. Biomechanical and physiological comparison of barefoot and two shod conditions in experienced barefoot runners. *J. Sports Med. Phys. Fitness* 6–13 (2009).

239. Boyer, E. R., Rooney, B. D. & Derrick, T. R. Rearfoot and midfoot or forefoot impacts in habitually shod runners. *Med. Sci. Sport. Exerc.* **46,** 1384–91 (2014).

240. Almonroeder, T., Willson, J. D. & Kernozek, T. W. The effect of foot strike pattern on achilles tendon load during running. *Ann. Biomed. Eng.* **41,** 1758–66 (2013).

241. Mussa-Ivaldi, F. A., Giszter, S. F. & Bizzi, E. Linear combinations of primitives in vertebrate motor control. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 7534–7538 (1994).

242. Tresch, M. C., Saltiel, P., D'Avella, A. & Bizzi, E. Coordination and localization in spinal motor systems. *Brain Res. Rev.* **40,** 66–79 (2002).

243. Saltiel, P., Wyler-Duda, K., D'Avella, A., Tresch, M. C. & Bizzi, E. Muscle synergies encoded within the spinal cord: evidence from focal intraspinal NMDA iontophoresis in the frog. *J. Neurophysiol.* **85,** 605–619 (2001).

244. Roh, J., Cheung, V. C.-K. & Bizzi, E. Modules in the brain stem and spinal cord underlying motor behaviors. *J. Neurophysiol.* **106,** 1363–1378 (2011).

245. Hart, C. B. & Giszter, S. F. Modular Premotor Drives and Unit Bursts as Primitives for Frog Motor Behaviors. *J. Neurosci.* **24,** 5269–5282 (2004).

246. Armstrong, D. M. The supraspinal control of mammalian locomotion. *J. Physiol.* **405,** 1–37 (1988).

247. Overduin, S. A., D'Avella, A., Carmena, J. M. & Bizzi, E. Microstimulation activates a handful of muscle synergies. *Neuron* **76,** 1071–7 (2012).

248. Ortiz-Rosario, A. & Adeli, H. Brain-computer interface technologies: from signal to action. *Rev. Neurosci.* **24,** 537–552 (2013).

249. Zhang, D., Huang, B., Wu, W. & Li, S. An Idle-State Detection Algorithm for SSVEP-Based Brain-Computer Interfaces Using a Maximum Evoked Response Spatial Filter. *Int. J. Neural Syst.* **25,** 1550030 (2015).

250. Burns, A., Adeli, H. & Buford, J. A. Brain-Computer Interface after Nervous System Injury. *Neurosci.* **20,** 639–651 (2014).

251. Ortiz-Rosario, A., Adeli, H. & Buford, J. A. Wavelet methodology to improve single unit isolation in primary motor cortex cells. *J. Neurosci. Methods* **246,** 106–118 (2015).

252. Cheung, V. C.-K. *et al.* Muscle synergy patterns as physiological markers of motor cortical damage. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 14652–14656 (2012).

253. Gizzi, L., Nielsen, J. F., Felici, F., Ivanenko, Y. P. & Farina, D. Impulses of activation but not motor modules are preserved in the locomotion of subacute stroke patients. *J. Neurophysiol.* **106,** 202–210 (2011).

254. Cheung, V. C.-K. *et al.* Stability of muscle synergies for voluntary actions after cortical stroke in humans. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 19563–8 (2009).

255. Ivanenko, Y. P. *et al.* Temporal Components of the Motor Patterns Expressed by the Human Spinal Cord Reflect Foot Kinematics. *J. Neurophysiol.* **90,** 3555–3565 (2003).

256. Steele, K. M., Tresch, M. C. & Perreault, E. J. The number and choice of muscles impact the results of muscle synergy analyses. *Front. Comput. Neurosci.* **7,** 105 (2013).

257. Hug, F., Turpin, N. A., Dorel, S. & Guével, A. Smoothing of electromyographic signals can influence the number of extracted muscle synergies. *Clin. Neurophysiol.* **123,** 1895–1896 (2012).

258. Cheung, V. C.-K. & Tresch, M. C. Non-negative matrix factorization algorithms modeling noise distributions within the exponential family. in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society.* **5,** 4990–3 (2005).

259. Liu, Y. *et al.* Three-Dimensional Innervation Zone Imaging from Multi-Channel Surface EMG Recordings. *Int. J. Neural Syst.* **25,** 1550024 (2015).

260. DeVita, P., Janshen, L., Rider, P., Solnik, S. & Hortobágyi, T. Muscle work is biased toward energy generation over dissipation in non-level running. *J. Biomech.* **41,** 3354–3359 (2008).

261. Cheung, V. C.-K., D'Avella, A. & Bizzi, E. Adjustments of Motor Pattern for Load Compensation Via Modulated Activations of Muscle Synergies During Natural Behaviors. *J. Neurophysiol.* **101,** 1235–1257 (2009).

262. Karamanidis, K., Arampatzis, A. & Brüggemann, G.-P. Reproducibility of

electromyography and ground reaction force during various running techniques. *Gait Posture* **19,** 115–123 (2004).

263. Endres, D. M., Chiovetto, E. & Giese, M. A. Model selection for the extraction of movement primitives. *Front. Comput. Neurosci.* **7,** 185 (2013).

264. Akaike, H. Factor analysis and AIC. *Psychometrika* **52,** 317–332 (1987).

265. Staudenmann, D., Potvin, J. R., Kingma, I., Stegeman, D. F. & van Dieën, J. H. Effects of EMG processing on biomechanical models of muscle joint systems: sensitivity of trunk muscle moments, spinal forces, and stability. *J. Biomech.* **40,** 900–9 (2007).

266. Potvin, J. R. & Brown, S. H. M. Less is more: high pass filtering, to remove up to 99% of the surface EMG signal power, improves EMG-based biceps brachii muscle force estimates. *J. Electromyogr. Kinesiol.* **14,** 389–99 (2004).

267. D'Avella, A., Saltiel, P. & Bizzi, E. Combinations of muscle synergies in the construction of a natural motor behavior. *Nat. Neurosci.* **6,** 300–8 (2003).

268. von Tscharner, V., Goepfert, B. & Nigg, B. M. Changes in EMG signals for the muscle tibialis anterior while running barefoot or with shoes resolved by non-linearly scaled wavelets. *J. Biomech.* **36,** 1169–1176 (2003).

269. Komi, P. V. in *The Encyclopaedia of sports medicine* (ed. Komi, P. V.) **III,** 184–202 (Blackwell Science Ltd, 1992).

270. Komi, P. V. Physiological and biomechanical correlates of muscle function: effects of muscle structure and stretch-shortening cycle on force and speed. *Exerc. Sport Sci. Rev.* **12,** 81–121 (1984).

271. Hägglund, M. *et al.* Optogenetic dissection reveals multiple rhythmogenic modules underlying locomotion. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 11589–11594 (2013).

272. Ivanenko, Y. P., Poppele, R. E. & Lacquaniti, F. Spinal Cord Maps of Spatiotemporal Alpha-Motoneuron Activation in Humans Walking at Different Speeds. *J.*

Neurophysiol. **95,** 602–618 (2006).

273. Monda, V. *et al.* Primary Motor Cortex Excitability in Karate Athletes: A Transcranial Magnetic Stimulation Study. *Front. Physiol.* **8,** 1–7 (2017).

274. Moscatelli, F. *et al.* Differences in corticospinal system activity and reaction response between karate athletes and non-athletes. *Neurol. Sci.* **37,** 1947–1953 (2016).

275. Moscatelli, F. *et al.* Functional Assessment of Corticospinal System Excitability in Karate Athletes. *PLoS One* **11,** e0159846 (2016).

276. Cappellini, G., Ivanenko, Y. P., Dominici, N., Poppele, R. E. & Lacquaniti, F. Migration of Motor Pool Activity in the Spinal Cord Reflects Body Mechanics in Human Locomotion. *J. Neurophysiol.* **104,** 3064–3073 (2010).

277. Ivanenko, Y. P., Cappellini, G., Poppele, R. E. & Lacquaniti, F. Spatiotemporal organization of α-motoneuron activity in the human spinal cord during different gaits and gait transitions. *Eur. J. Neurosci.* **27,** 3351–3368 (2008).

278. La Scaleia, V., Ivanenko, Y. P., Zelik, K. E. & Lacquaniti, F. Spinal motor outputs during step-to-step transitions of diverse human gaits. *Front. Hum. Neurosci.* **8,** 305 (2014).

279. Cappellini, G. *et al.* Immature spinal locomotor output in children with cerebral palsy. *Front. Physiol.* **7,** 1–21 (2016).

280. Kendall, F. P., McCreary, E. K., Provance, P. G., Rodgers, M. M. & Romani, W. A. *Muscles: Testing and Function, with Posture and Pain*. (Lippincott Williams & Wilkins, 2005).

281. De Wit, B., De Clercq, D. & Aerts, P. Biomechanical analysis of the stance phase during barefoot and shod running. *J. Biomech.* **33,** 269–278 (2000).

282. Lee, S. S. M. & Piazza, S. J. Built for speed: musculoskeletal structure and sprinting ability. *J. Exp. Biol.* **212,** 3700–7 (2009).

283. Ekizos, A., Santuz, A. & Arampatzis, A. Transition from shod to barefoot alters dynamic stability during running. *Gait Posture* **56,** 31–36 (2017).

284. Kelly, L. A., Lichtwark, G. A., Farris, D. J. & Cresswell, A. Shoes alter the spring-like function of the human foot during running. *J. R. Soc. Interface* **13,** 20160174 (2016).

285. Pearson, K. G., Ekeberg, Ö. & Büschges, A. Assessing sensory function in locomotor systems using neuro-mechanical simulations. *Trends Neurosci.* **29,** 625–631 (2006).

286. IEEE Standards Association. *ISO/IEC/IEEE 24765:2010 Systems and software engineering - Vocabulary*. *Iso/Iec/Ieee 24765:2010* **25021,** (Institute of Electrical and Electronics Engineers, Inc., 2010).

287. Kitano, H. Biological robustness. *Nat. Rev. Genet.* **5,** 826–837 (2004).

288. Maiwald, C., Sterzing, T., Mayer, T. A. & Milani, T. L. Detecting foot-to-ground contact from kinematic data in running. *Footwear Sci.* **1,** 111–118 (2009).

289. Kantz, H. & Schreiber, T. *Nonlinear Time Series Analysis*. **7,** (Cambridge University Press, 2004).

290. Schreiber, T. Detecting and Analyzing Nonstationarity in a Time Series Using Nonlinear Cross Predictions. *Phys. Rev. Lett.* **78,** 843–846 (1997).

291. Bruijn, S. M., van Dieën, J. H., Meijer, O. G. & Beek, P. J. Statistical precision and sensitivity of measures of dynamic gait stability. *J. Neurosci. Methods* **178,** 327–333 (2009).

292. Kang, H. G. & Dingwell, J. B. Intra-session reliability of local dynamic stability of walking. *Gait Posture* **24,** 386–90 (2006).

293. Packard, N. H., Crutchfield, J. P., Farmer, J. D. & Shaw, R. S. Geometry from a time series. *Physical Review Letters* **45,** 712–716 (1980).

294. Takens, F. in *Dynamical Systems and Turbulence, Warwick 1980* (eds. Rand, D. & Young, L.-S.) 366–381 (Springer Berlin Heidelberg, 1981). doi:10.1007/BFb0091924

295. Fraser, A. M. & Swinney, H. L. Independent coordinates for strange attractors from mutual information. *Physical Review A* **33,** 1134–1140 (1986).

296. Kennel, M. B., Brown, R. & Abarbanel, H. D. I. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* **45,** 3403–3411 (1992).

297. Bradley, E. & Kantz, H. Nonlinear time-series analysis revisited. *Chaos* **25,** 097610 (2015).

298. Kugiumtzis, D. State space reconstruction parameters in the analysis of chaotic time series--the role of the time window length. *Phys. D* **95,** 13–28 (1996).

299. Rosenstein, M. T., Collins, J. J. & De Luca, C. J. A practical method for calculating largest Lyapunov exponents from small data sets. *Phys. D* **65,** 117–134 (1993).

300. Dingwell, J. B. & Cusumano, J. P. Nonlinear time series analysis of normal and pathological human walking. *Chaos* **10,** 848–863 (2000).

301. Terrier, P. & Dériaz, O. Non-linear dynamics of human locomotion: Effects of rhythmic auditory cueing on local dynamic stability. *Front. Physiol.* **4 SEP,** 1–13 (2013).

302. Hoogkamer, W. *et al.* Toward new sensitive measures to evaluate gait stability in focal cerebellar lesion patients. *Gait Posture* **41,** 592–596 (2015).

303. Reynard, F., Vuadens, P., Deriaz, O. & Terrier, P. Could local dynamic stability serve as an early predictor of falls in patients with moderate neurological gait disorders? A reliability and comparison study in healthy individuals and in patients with paresis of the lower extremities. *PLoS One* **9,** (2014).

304. McAndrew, P. M., Dingwell, J. B. & Wilken, J. M. Walking variability during

continuous pseudo-random oscillations of the support surface and visual field. *J. Biomech.* **43,** 1470–1475 (2010).

305. Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* **47,** 381–391 (1954).

306. Meghdadi, A. H. On robustness of evolutionary fuzzy control systems. in *IEEE Annual Meeting of the Fuzzy Information, 2004. Processing NAFIPS '04.* 254–258 Vol.1 (IEEE, 2004). doi:10.1109/NAFIPS.2004.1336287

307. More, H. L. *et al.* Sensorimotor responsiveness and resolution in the giraffe. *J. Exp. Biol.* **216,** 1003–1011 (2013).

308. Hamill, J. & Gruber, A. H. Is changing footstrike pattern beneficial to runners? *J. Sport Heal. Sci.* **6,** 146–153 (2017).

309. Davis, I. S., Rice, H. M. & Wearing, S. C. Why forefoot striking in minimal shoes might positively change the course of running injuries. *J. Sport Heal. Sci.* **6,** 154–161 (2017).

310. Squadrone, R., Rodano, R., Hamill, J. & Preatoni, E. Acute effect of different minimalist shoes on foot strike pattern and kinematics in rearfoot strikers during running. *J. Sports Sci.* **33,** 1196–204 (2015).

311. Gillinov, S. M., Laux, S., Kuivila, T., Hass, D. & Joy, S. M. Effect of Minimalist Footwear on Running Efficiency: A Randomized Crossover Trial. *Sports Health* **7,** 256–60 (2015).

312. Santuz, A., Ekizos, A. & Arampatzis, A. Effects of exercise-induced changes in muscle mechanical advantage on energy cost and modular organisation of running. in *22nd Annual Congress of the European College of Sport Science* (2017).

313. Ekizos, A., Santuz, A. & Arampatzis, A. Short- and long-term effects of altered point of ground reaction force application on human running energetics. *J. Exp. Biol.* jeb.176719 (2018). doi:10.1242/jeb.176719

314. Boisgontier, M. P., Olivier, I., Chenu, O. & Nougier, V. Presbypropria: the effects of physiological ageing on proprioceptive control. *Age (Omaha).* **34,** 1179–1194 (2012).

315. Robles-De-La-Torre, G. The Importance of the Sense of Touch in Virtual and Real Environments. *IEEE Multimed.* **13,** 24–30 (2006).

316. Suzuki, T., Kinugasa, R. & Fukashiro, S. Activation of plantar flexor muscles is constrained by multiple muscle synergies rather than joint torques. *PLoS One* **12,** 1–16 (2017).

317. Cruz Ruiz, A. L., Pontonnier, C. & Dumont, G. Low-Dimensional Motor Control Representations in Throwing Motions. *Appl. Bionics Biomech.* **2017,** 1–19 (2017).

318. Reinkensmeyer, D. J., Emken, J. L. & Cramer, S. C. Robotics, Motor Learning, and Neurologic Recovery. *Annu. Rev. Biomed. Eng.* **6,** 497–525 (2004).

319. Federici, S., Meloni, F., Bracalenti, M. & De Filippis, M. L. The effectiveness of powered, active lower limb exoskeletons in neurorehabilitation: A systematic review. *NeuroRehabilitation* **37,** 321–340 (2015).

320. Jacobs, D. A., Koller, J. R., Steele, K. M. & Ferris, D. P. Motor modules during adaptation to walking in a powered ankle exoskeleton. *J. Neuroeng. Rehabil.* **15,** 2 (2018).

321. Alibeji, N. A., Kirsch, N. A. & Sharma, N. A Muscle Synergy-Inspired Adaptive Control Scheme for a Hybrid Walking Neuroprosthesis. *Front. Bioeng. Biotechnol.* **3,** 1–13 (2015).

322. Sylos-Labini, F. *et al.* EMG patterns during assisted walking in the exoskeleton. *Front. Hum. Neurosci.* **8,** 1–12 (2014).

323. Patane, F., Rossi, S., Del Sette, F., Taborri, J. & Cappa, P. WAKE-Up Exoskeleton to Assist Children With Cerebral Palsy: Design and Preliminary Evaluation in Level Walking. *IEEE Trans. Neural Syst. Rehabil. Eng.* **25,** 906–916 (2017).

324. Rossi, S. *et al.* Feasibility Study of a Wearable Exoskeleton for Children: Is the Gait Altered by Adding Masses on Lower Limbs? *PLoS One* **8,** e73139 (2013).

Acknowledgements

Not a single page of this work would have been written without the constant and sincere support of my two supervisors Diamantis (Prof. Adamantios Arampatzis) and Bill (Prof. Vasilios Baltzopoulos). They were always there for any of my scientific or personal need and helped me to see academia in a perspective I never considered before.

The years in the department flew by in a blink of an eye also thanks to my colleagues and the wonderful working environment they are part of. Both my personal and scientific growth found fertile ground in the hours spent talking about science and life decisions.

The almost 300 persons who participated in our studies gave me the wonderful possibility to get to know new people and have a deeper understanding of German society and culture. They have been essential part of this journey.

My friends continuously supported my choices and pushed me to do better. By means of some magic that nothing has to do with science, they made the physical distance easier to bear on a daily basis.

There are not enough words nor pages to manifest the gratitude and love I feel towards my family, which is possibly the main reason why I am doing this. Especially my parents and my uncle always made me feel like I did the right choices and their being proud of what I do is a gift every day. Katrina, with her unconditional love, has been, is and will be the immovable rock of my life and that does not need further explanations.

Declaration

Ich erkläre, dass ich die vorliegende Dissertation selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe. Alle Zitate sowie sinngemäße wörtliche Wiedergaben, die anderen Werken entnommen wurden, sind unter Angabe der Quelle kenntlich gemacht.

I hereby declare that I have completed this doctoral thesis independently. All the sources and aids I used to support my work were explicitly declared and listed or, where specified, literally quoted.